

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 18:11:59 ; Search time 5268 Seconds
(without alignments)
11082.587 Million cell updates/sec

Title: US-10-650-369-21
Perfect score: 1347
Sequence: 1 atgaaaaaaataacaggat.....ttgcaaaaatcgtaataaa 1347

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1347	100.0	1347	6	AR437352	Sequence
2	1347	100.0	1347	6	AX367176	Sequence
3	817.4	60.7	1011	1	AF375662	Streptoco
4	815.8	60.6	1011	6	AR437347	Sequence
5	815.8	60.6	1011	6	AX355986	Sequence
6	815.8	60.6	1011	6	AX367166	Sequence
7	782.2	58.1	1158	1	AB088216	Streptoco
8	782.2	58.1	1158	1	AB088217	Streptoco
9	782.2	58.1	2431	1	AB088214	Streptoco
10	780.6	58.0	1104	1	AB088215	Streptoco
11	780.6	58.0	1125	6	AR135510	Sequence
12	780.6	58.0	1156	1	STRPLR	M95569 Streptococ
13	780.6	58.0	2326	1	AF520813	Streptoco
14	780.6	58.0	10053	1	AE006494	Streptoco
15	780.6	58.0	50029	1	AE014140	Streptoco
16	780.6	58.0	311600	1	AP005141	Streptoco
17	779	57.8	12001	1	AE009973	Streptoco
18	757.6	56.2	1260	1	SEGAPCGEN	X97788 S.equisimil
19	757.6	56.2	2334	1	SEGAPLPP	Y12602 Streptococ
20	702.2	52.1	1011	6	AR437348	Sequence
21	702.2	52.1	1011	6	AX355988	Sequence
22	702.2	52.1	1011	6	AX367168	Sequence
23	700.6	52.0	1011	6	AX607902	Sequence
24	700.6	52.0	12960	6	AX602161	Sequence
25	700.6	52.0	122050	1	SAG766853	Streptoco
26	699	51.9	1011	1	AF338416	Streptoco
27	699	51.9	21954	1	AE014272	Streptoco
28	694.2	51.5	1092	1	AB110908	Streptoco
29	686.2	50.9	1012	1	AY167026	Streptoco
30	683	50.7	7034	2	SPNEU1925	Streptoco
31	678.2	50.3	1008	1	SPN505822	Streptoco
32	678.2	50.3	1011	1	AF421899	Streptoco
33	678.2	50.3	1011	1	AF421902	Streptoco
34	678.2	50.3	1011	6	AR437351	Sequence
35	678.2	50.3	1011	6	AX355994	Sequence
36	678.2	50.3	1011	6	AX367174	Sequence
37	678.2	50.3	7571	6	AR218770	Sequence
38	678.2	50.3	7571	6	BD003682	Polynucle
39	678.2	50.3	10145	1	AE007490	Streptoco
40	678.2	50.3	12139	1	AE008547	Streptoco
41	678.2	50.3	34980	6	AX571765	Sequence
42	678.2	50.3	34980	6	AX571766	Sequence
43	676.6	50.2	1011	1	AF421900	Streptoco
44	676.6	50.2	1011	6	AR437349	Sequence
45	676.6	50.2	1011	6	AX355990	Sequence

ALIGNMENTS

RESULT 1
AR437352
LOCUS AR437352 1347 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 21 from patent US 6660270.
ACCESSION AR437352
VERSION AR437352.1 GI:40202283
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1347)
AUTHORS Potter,A.A., Perez-Casal,J. and Fontaine,M.
TITLE Immunization of dairy cattle with chimeric GapC protein against Streptococcus infection
JOURNAL Patent: US 6660270-A 21 09-DEC-2003;

Pred. No. is the number of results predicted by chance to have a

FEATURES		Location/Qualifiers	
source		1..1347	
		/organism="unknown"	
		/mol_type="genomic DNA"	
ORIGIN			
Query Match		100.0%;	Score 1347; DB 6; Length 1347;
Best Local Similarity		100.0%;	Pred. No. 1.4e-261;
Matches 1347; Conservative		0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	ATGAAAAAATAACAGGGATTATTTTATTGCTCTCTTGCAGTCATTATTCTGTGTCATGC	60
Db	1	ATGAAAAAATAACAGGGATTATTTTATTGCTCTCTTGCAGTCATTATTCTGTGTCATGC	60
QY	61	CAGGCAAACTACGGATCCGGTATGGTAGTTAAAGTTGGAAGTGTGTAACGGTTTCGGTCGTATC	120
Db	61	CAGGCAAACTACGGATCCGGTATGGTAGTTAAAGTTGGAAGTGTGTAACGGTTTCGGTCGTATC	120
QY	121	GGACGCTCTTGCATTCCGTCGTATTCAAAATGTTGAAGTGTGTAAGTAACTCGTATCAAC	180
Db	121	GGACGCTCTTGCATTCCGTCGTATTCAAAATGTTGAAGTGTGTAAGTAACTCGTATCAAC	180
QY	181	GACCTTACAGATCCAAACATGCTTGCACACTTGTGAAATACGATACAACTCAAGGACGT	240
Db	181	GACCTTACAGATCCAAACATGCTTGCACACTTGTGAAATACGATACAACTCAAGGACGT	240
QY	241	TTTGACGGAACCTGTTGAAGTTAAAGAGGTGGAATTTGAAGTAAACGGAACCTTCATCAA	300
Db	241	TTTGACGGAACCTGTTGAAGTTAAAGAGGTGGAATTTGAAGTAAACGGAACCTTCATCAA	300
QY	301	GTTTCTGCTGAACGTCATCCAGAAAAACATCGACTGGCAACTGACGGTGTGAAATCGTT	360
Db	301	GTTTCTGCTGAACGTCATCCAGAAAAACATCGACTGGCAACTGACGGTGTGAAATCGTT	360
QY	361	CTGGAAGCACTCGAGGTCATGTAAGTTAAAGATGGTGGATTGACGTTAAACGGAATA	420
Db	361	CTGGAAGCACTCGAGGTCATGTAAGTTAAAGATGGTGGATTGACGTTAAACGGAATA	420
QY	421	TTTCAATAAGTTTCTGCTGAAAAAGATCCAGAACAAATTGACTGGCAACTGACGGTGT	480
Db	421	TTTCAATAAGTTTCTGCTGAAAAAGATCCAGAACAAATTGACTGGCAACTGACGGTGT	480
QY	481	GAAATCGTTCTTGAATCGATGGTACTGTTGAGTTAAAGAGGTGGAATTCGAAGTTAAC	540
Db	481	GAAATCGTTCTTGAATCGATGGTACTGTTGAGTTAAAGAGGTGGAATTCGAAGTTAAC	540
QY	541	GGTCAATTTGTTAAAGTTTCTGCTGAACCGCAACCAAGCAATTTGACTGGGCTACTGAT	600
Db	541	GGTCAATTTGTTAAAGTTTCTGCTGAACCGCAACCAAGCAATTTGACTGGGCTACTGAT	600
QY	601	GGCGTAGAAATCGTTTCTGAAGCAACTAGTTTCTTGTCTAAAAAGAGCTGCTGAAAA	660
Db	601	GGCGTAGAAATCGTTTCTGAAGCAACTAGTTTCTTGTCTAAAAAGAGCTGCTGAAAA	660
QY	661	CACCTACATGCTAACGGTGTCTAAAAAGTTTGTATCACAGCTCCTGTTGGAACGACGTT	720
Db	661	CACCTACATGCTAACGGTGTCTAAAAAGTTTGTATCACAGCTCCTGTTGGAACGACGTT	720
QY	721	AAAAACAGTTGTTTCAACACTAACCAACGACATCTTGAACGGTACTGAAACAGTTATCTCA	780
Db	721	AAAAACAGTTGTTTCAACACTAACCAACGACATCTTGAACGGTACTGAAACAGTTATCTCA	780
QY	781	GGTGCTTCATGTACTACAAACTGTTTAGCTCCTATGGCTAAAGCTCTTACGATGCATTT	840
Db	781	GGTGCTTCATGTACTACAAACTGTTTAGCTCCTATGGCTAAAGCTCTTACGATGCATTT	840
QY	841	GGTATCCAAAAAGGCTTATGACTACAAATCCAGCTTATACCTGGTGACCAAAATGATCCTT	900
Db	841	GGTATCCAAAAAGGCTTATGACTACAAATCCAGCTTATACCTGGTGACCAAAATGATCCTT	900
QY	901	GACGGACCAACACCGTGGTGGTGACCTTCGTGCTGCTCGTGGTGGTGCACAAACATTGTT	960
Db	901	GACGGACCAACACCGTGGTGGTGACCTTCGTGCTGCTCGTGGTGGTGCACAAACATTGTT	960

QY	961	CCTAACTCAACTGGTGTCTGCTAAAGCTATCGTCTTGTATTATCCAGAAATGAATGGTAAA	1020
Db	961	CCTAACTCAACTGGTGTCTGCTAAAGCTATCGTCTTGTATTATCCAGAAATGAATGGTAAA	1020
QY	1021	CTTGATGGTGTCTGCAACACGTTTCTCTGTTCCAACTGGATCAGTAACAGTTGGTTGTA	1080
Db	1021	CTTGATGGTGTCTGCAACACGTTTCTCTGTTCCAACTGGATCAGTAACAGTTGGTTGTA	1080
QY	1081	ACTCTTGATAAAAAACGTTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAAC	1140
Db	1081	ACTCTTGATAAAAAACGTTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAAC	1140
QY	1141	GACAGTTTTCGGTTACACTGAAGATCCAAATTTCTTCCAGATATCGTAGCGTGTATAC	1200
Db	1141	GACAGTTTTCGGTTACACTGAAGATCCAAATTTCTTCCAGATATCGTAGCGTGTATAC	1200
QY	1201	GGTTCATTGTTTGACGCAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTT	1260
Db	1201	GGTTCATTGTTTGACGCAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTT	1260
QY	1261	AAAGTTGTATCATGTTATGACAATGAAATGTTCTTACACTGCTCAACTTGTTCGTACACTT	1320
Db	1261	AAAGTTGTATCATGTTATGACAATGAAATGTTCTTACACTGCTCAACTTGTTCGTACACTT	1320
QY	1321	GAGTATTTTGCAAAAATCGCTAAATAA	1347
Db	1321	GAGTATTTTGCAAAAATCGCTAAATAA	1347

RESULT 2			
AX3671176			
LOCUS	AX3671176	1347 bp	DNA linear PAT 16-FEB-2002
DEFINITION	Sequence 21 from Patent WO0196379.		
ACCESSION	AX3671176		
VERSION	AX3671176.1	GI:18855374	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1		
AUTHORS	Potter,A.A., Perez-Casal,J. and Fontaine,M.		
TITLE	Immunization of dairy cattle with chimeric gapC protein against		
JOURNAL	Streptococcus infection		
	Patent: WO 0196379-A 21 20-DEC-2001;		
	The University of Saskatchewan (CA)		
FEATURES	Location/Qualifiers		
source	1..1347		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Primer Gap4 chimeric GapC protein"		
CDS	1..1347		
	/note="unnamed protein product"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAD23843.1"		
	/db_xref="GI:18855375"		
	/db_xref="REMTREMBL:CAD23843"		
	/translation="MKITGIILLLAVIILSACQANYGSMVVKVINGFGRIGRLA		
	FRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEKVEGGFEVNGNFIKVS		
	AERDPENIDWATDGEIVLEALEGTVBVKGGDVGKFIKVSABKDPQIDWATDGV		
	EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLEATSFFAKKEAA		
	EKHLHANGAKVWITAPGNDVKTWVFNTNHDILDGTETVSGASCTTNCLAPMAKAL		
	HDAFGIQGLMTTHAYTGDMILDPHRRGGDLRRARAGAAIVPNSTGAAKAIGLVI		
	PELNGKLDGAAQRVPVPTGVTELVTLTKNVSVDIINAAKKAASNDSTFGYTEDPIVS		
	SDIVGVSYGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK"		

Query Match	100.0%;	Score 1347; DB 6; Length 1347;
Best Local Similarity	100.0%;	Pred. No. 1.4e-261;
Matches 1347; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAATAACAGGATTTATTTATTTGCTTCTTGCAGTCATTATTCTGTCTGCATGC 60
Db 1 ATGAAAAAATAACAGGATTTATTTATTTGCTTCTTGCAGTCATTATTCTGTCTGCATGC 60
QY 61 CAGGCAAACTACCGGATCCGGTATGGTAGTTAAAGTTGGTATTAACGGTTTCGGTCTGATC 120
Db 61 CAGGCAAACTACCGGATCCGGTATGGTAGTTAAAGTTGGTATTAACGGTTTCGGTCTGATC 120
QY 121 GGACGCTCTGCATTCCTGTCGATTAACAAATGTTGAAGGTTGAAGTAACTCGTATCAAC 180
Db 121 GGACGCTCTGCATTCCTGTCGATTAACAAATGTTGAAGGTTGAAGTAACTCGTATCAAC 180
QY 181 GACCTTACAGATCCAAACATGCTTGCACACTTGTGTAATACGATACAACTCAAGGACGT 240
Db 181 GACCTTACAGATCCAAACATGCTTGCACACTTGTGTAATACGATACAACTCAAGGACGT 240
QY 241 TTTGACGGAACCTGTTGAAGTTAAAGAGGTGGATTGTAAGTAAACGGAACTTTCATCAAA 300
Db 241 TTTGACGGAACCTGTTGAAGTTAAAGAGGTGGATTGTAAGTAAACGGAACTTTCATCAAA 300
QY 301 GTTCTGCTGAACGATGATCCAGAAACATCGACTGGGCAACTGACGGTGTGTAATCGTT 360
Db 301 GTTCTGCTGAACGATGATCCAGAAACATCGACTGGGCAACTGACGGTGTGTAATCGTT 360
QY 361 CTGGAAGCACTCGAGGTTACTGTAGAAAGTTAAAGATGTTGATTTGACGTTAAACGGAAAA 420
Db 361 CTGGAAGCACTCGAGGTTACTGTAGAAAGTTAAAGATGTTGATTTGACGTTAAACGGAAAA 420
QY 421 TTCATTAAGTTTCTGCTGAAAGATCCAGAAACAAATGACTGGGCAACTGACGGTGT 480
Db 421 TTCATTAAGTTTCTGCTGAAAGATCCAGAAACAAATGACTGGGCAACTGACGGTGT 480
QY 481 GAAATCGTTCTTGAATCGATGTTGAAAGTTAAAGATGTTGATTTGACGTTAAACGGAAAA 540
Db 481 GAAATCGTTCTTGAATCGATGTTGAAAGTTAAAGATGTTGATTTGACGTTAAACGGAAAA 540
QY 541 GGTCAATTTGTTAAAGTTTCTGCTGAACGCAACCGCAACATGACTGGGCTACTGAT 600
Db 541 GGTCAATTTGTTAAAGTTTCTGCTGAACGCAACCGCAACATGACTGGGCTACTGAT 600
QY 601 GCGTAGAAATCGTTCTTGAAGCAACTAGTTTCTTGTGCTAAAGAAAGCTGCTGAAAAA 660
Db 601 GCGTAGAAATCGTTCTTGAAGCAACTAGTTTCTTGTGCTAAAGAAAGCTGCTGAAAAA 660
QY 661 CACTTACATGCTAACGGTGTGCTAAAGAAAGTTGTTATACAGCTCCTGGTGAACACGCTT 720
Db 661 CACTTACATGCTAACGGTGTGCTAAAGAAAGTTGTTATACAGCTCCTGGTGAACACGCTT 720
QY 721 AAAACAGTTGTTTCAACACTAACCCAGCACTTCTGACGGTACTGAAACAGTTATCTCA 780
Db 721 AAAACAGTTGTTTCAACACTAACCCAGCACTTCTGACGGTACTGAAACAGTTATCTCA 780
QY 781 GGTGCTTCATGTAATAAAGTTTGTAGCTCCTATGGCTAAAGCTCTTACGATGCAATTT 840
Db 781 GGTGCTTCATGTAATAAAGTTTGTAGCTCCTATGGCTAAAGCTCTTACGATGCAATTT 840
QY 841 GGTATCCAAAAAGGCTTATGACTACAAATCCAGCTTATCTGTTGACCAATGATCCTT 900
Db 841 GGTATCCAAAAAGGCTTATGACTACAAATCCAGCTTATCTGTTGACCAATGATCCTT 900
QY 901 GACGGACCAACCGTGGTGGTACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 960
Db 901 GACGGACCAACCGTGGTGGTACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 960
QY 961 CCTAACTCAACTGGTGGTGGTAAAGCTATCGGCTCTGTTATCCAGAAATGATGTTAA 1020
Db 961 CCTAACTCAACTGGTGGTGGTAAAGCTATCGGCTCTGTTATCCAGAAATGATGTTAA 1020
QY 1021 CTTGATGGTGTGCACAACTGTTCTGTTCCAACTGGATCAGTAACAGTTGTTGTA 1080
Db 1021 CTTGATGGTGTGCACAACTGTTCTGTTCCAACTGGATCAGTAACAGTTGTTGTA 1080
QY 1081 ACTCTTGATAAAAAACGTTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTCTTCAAC 1140

Db 1081 ACTCTTGATAAAAAACGTTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTCTTCAAC 1140
QY 1141 GACAGTTTTCGGTTTACACTGAAGATCCAAATGTTTCTTCCAGATATCGTAGCGTGTACATC 1200
Db 1141 GACAGTTTTCGGTTTACACTGAAGATCCAAATGTTTCTTCCAGATATCGTAGCGTGTACATC 1200
QY 1201 GGTTCATTGTTTACGCAACTCAACTAAAGTTATGGAAGTTGACGGAATCAATTTGGTT 1260
Db 1201 GGTTCATTGTTTACGCAACTCAACTAAAGTTATGGAAGTTGACGGAATCAATTTGGTT 1260
QY 1261 AAAGTTGTTATCATGATGACAAATGAAATGTTTCTTACACTGCTCAACTGTTCTGACACTT 1320
Db 1261 AAAGTTGTTATCATGATGACAAATGAAATGTTTCTTACACTGCTCAACTGTTCTGACACTT 1320
QY 1321 GAGTATTTTGCAAAAATCGCTAAATAA 1347
Db 1321 GAGTATTTTGCAAAAATCGCTAAATAA 1347
RESULT 3
AF375662 1011 bp DNA linear BCT 03-MAY-2003
LOCUS Streptococcus dysgalactiae glycerolaldehyde 3-P dehydrogenase (gapC)
DEFINITION gene, complete cds.
ACCESSION AF375662
VERSION AF375662.1 GI:30348860
KEYWORDS Streptococcus dysgalactiae
SOURCE Streptococcus dysgalactiae
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 1011)
AUTHORS Perez-Casal,J.F., Bolton,A., Song,X.-M., Willson,P. and Potter,A.A.
TITLE Use of the surface proteins GapC and Mig of Streptococcus
dysgalactiae as protective antigens against mastitis in
non-lactating cows
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1011)
AUTHORS Perez-Casal,J.F., Bolton,A., Song,X.-M., Willson,P. and Potter,A.A.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2001) VIDO, University of Saskatchewan, 120
Veterinary Rd, Saskatoon, SK S7N 5E3, Canada
FEATURES
source Location/Qualifiers
1..1011
/organism="Streptococcus dysgalactiae"
/mol_type="genomic DNA"
/db_xref="taxon:1334"
1..1011
/gene="gapC"
1..1011
/note="GapC"
/codon_start=1
/transl_table=11
/product="glycerolaldehyde 3-P dehydrogenase"
/protein_id="AAP31408.1"
/db_xref="GI:30348861"
/translation="MVVKVINGFGRIGRIARFRIQNVGEVVEVTRINDLTDNMLAHL
LKYDTTQGRFDGTVVEKGGFEVNGFNFKVSAERDPENIDWATDGVIEATGFFAK
KEAAEKHLHANGAKVVTAPGNDVKTVPNTNHDILDGTEVTSAGSCTNCLAPM
AKALHDAFGIQGLMTTHAYTGDQMLDGPHRGDLRRARAGAAANIVNSTGAAKAI
GLVIPELNGKLDGAAQRPVPTGVSVELVVTLDKNVSVDEINAAKAAASNDVSFGYTED
PIVSSDIVGVSVGSLFDATQTQKMEVDGSQLVKVSVSWYDNEMSYTAQLVRLTYEYFAKI
AK"
ORIGIN
Query Match 60.7%; Score 817.4; DB 1; Length 1011;
Best Local Similarity 97.5%; Pred. No. 1.1e-154;
Matches 830; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 497 TCATGGTACTGTTGAAGTTAAAGAGGTGGATTCGAAAGTTAACGGTCAATTTGTTAAAG 556

[illegible]

```
Db 161 TTGACGGAAGTGTGAAGTTAAAGAGGTGATTGAAGTAAACGGAAACTTCATCAAAG 220
QY 557 TTCTGTCTGAACGCGAACACAGCAACATGACTGGGCTACTGATGGCGTAGAAATCGTTC 616
Db 221 TTCTGTCTGAACGCGAACACATGACTGGGCAACTGACGGTGTGAAATCGTTC 280
QY 617 TTGAAGCAACTAGTTTCTTGCTAAAGAGCTGCTGAAACACATTCATGCTAAAG 676
Db 281 TGGAAAGCAACTGGTTCTTGCTAAAGAGAGCTGCTGAAACACATTCATGCTAAAG 340
QY 677 GTGCTAAAGAAAGTTGTTATACAGCTCCTGGTGGAAACGACGTTAAACAGTTGTTTCA 736
Db 341 GTGCTAAAGAAAGTTGTTATACAGCTCCTGGTGGAAACGACGTTAAACAGTTGTTTCA 400
QY 737 ACACATAACCAACGACATTTCTTGACGGTACTGAACAGTTATCTCAGGTGCTTCACTA 796
Db 401 ACACATAACCAACGACATTTCTTGACGGTACTGAACAGTTATCTCAGGTGCTTCACTA 460
QY 797 CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTACAGATGATCTTGGTATCCAAAAGGTC 856
Db 461 CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTACAGATGATCTTGGTATCCAAAAGGTC 520
QY 857 TTATGACTACAATCCACGCTTATCTGGTGACAAATGATCCTTGACGGACACACCGTG 916
Db 521 TTATGACTACAATCCACGCTTATCTGGTGACAAATGATCCTTGACGGACACACCGTG 580
QY 917 GTGGTGACCTTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 581 GTGGTGACCTTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
QY 977 CTGCTAAAGCTATCGGCTTGTGTTATCCAGAAATGAATGGTAACTGATGGTGTGCTGCAC 1036
Db 641 CTGCTAAAGCTATCGGCTTGTGTTATCCAGAAATGAATGGTAACTGATGGTGTGCTGCAC 700
QY 1037 AACGTGTTCTGTTCCAACTGGATCAGTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1096
Db 701 AACGTGTTCTGTTCCAACTGGATCAGTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 760
QY 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACACGACAGTTTCGGTTACA 1156
Db 761 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACACGACAGTTTCGGTTACA 820
QY 1157 CTGAAGATCCAAATGTTTCTTCAGATATCTGAGCGGTGTCATACGTTTATTGTTGACG 1216
Db 821 CTGAAGATCCAAATGTTTCTTCAGATATCTGAGCGGTGTCATACGTTTATTGTTGACG 880
QY 1217 CAACCTCAAACTAAAGTTATGGAAGTTGACGATCAACAATGGTTAAAGTTGATCATGGT 1276
Db 881 CAACCTCAAACTAAAGTTATGGAAGTTGACGATCAACAATGGTTAAAGTTGATCATGGT 940
QY 1277 ATGACAAATGAATGTCTTACACTGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1336
Db 941 ATGACAAATGAATGTCTTACACTGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1000
QY 1337 TCGCTAAATAA 1347
Db 1001 TCGCTAAATAA 1011

RESULT 7
AB088216
LOCUS
DEFINITION
Streptococcus pyogenes naplr gene for nephritis-associated plasmin
receptor, complete cds, strain:M type 12.
ACCESSION
AB088216
VERSION
AB088216.1 GI:21886629
KEYWORDS
Streptococcus pyogenes
Streptococcus pyogenes
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
SOURCE
Yoshizawa,N., Yamakami,K. and Fujino,M.
ORGANISM
REFERENCE
1
```

```
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
-35_signal
-10_signal
RBS
gene
CDS
/organism="Streptococcus pyogenes"
/mol_type="genomic DNA"
/strain="M type 12"
/db_xref="taxon:1314"
13..18
35..40
100..106
117..1127
/gene="naplr"
117..1127
/gene="naplr"
/EC_number="1.2.1.12"
/function="glyceraldehyde-3-phosphate dehydrogenase;
plasmin binding activity"
/note="glyceraldehyde-3-phosphate dehydrogenase (GAPDH);
plasmin binding protein; nephritogenic antigen"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="nephritis-associated plasmin receptor"
/protein_id="BAC05496.1"
/db_xref="GI:21886630"
/translation="MVVKVINGFGRIGRLAFRRIQNIEGVETRIINDLTDPNMLAHL
LKDYDTQGRFDGTVKEGVEGNFVKVSAERDPENIDWTDGVEIVLEATGFFAK
KEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTEVISGASCTTNCLAPM
AKALHDAFGIQKGLMTTIHAYTDQMLDGPGRGDLRRARAGAAIVPNSTGAAKAI
GLVIPELNGKLDGAAQRPVPTGVTSLVTLVDKNVSDVDEINAAKASNDSTFGYTED
PIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVWSYDNEVSYTAQLVRLTEYFAKI
AK"
```

ORIGIN

```
Query Match 58.1%; Score 782.2; DB 1; Length 1158;
Best Local Similarity 94.9%; Pred. No. 1.4e-147;
Matches 808; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 497 TCGATGGTACTGTTGAAGTTAAAGAGGTGATTCGAAGTTAACGGTCAATTTGTTAAAG 556
Db 277 TTGATGGAACAGTTGAAGTTAAAGAGGTGATTTGAAGTAAACGGAACCTTCATTAAG 336
QY 557 TTTCTGCTGAACGCGGAACACAGCAACATTCAGCTGGGCTACTGATGGCGTAGAAATCGTTC 616
Db 337 TTTCTGCTGAACGCTGATCCAGAAACATCGACTGGGCACTGATGGGTTGAAATCGTTC 396
QY 617 TTGAAGCAACTAGTTTCTTTGCTAAAGAAAGAGCTGCTGAAACACACTTACATGCTAACG 676
Db 397 TTGAAGCAACTAGTTTCTTTGCTAAAGAAAGAGAGCAGCTGAAACACACTTACATGCTAACG 456
QY 677 GTGCTAAAGAAAGTTGTTATCACAGCTCCTGGTGGAAACGACGTTAAACAGTTGTTTCA 736
Db 457 GTGCTAAAGAAAGTTGTTATCACAGCTCCTGGTGGAAACGATGTTAAACAGTTGTTTCA 516
QY 737 ACACATAACCAACGACATTTCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCACTA 796
Db 517 ACACATAACCAACGACATTTCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCACTA 576
QY 797 CAAACTGTTTAGCTCCTATGGCTAAAGCTTTCACGATGCTATTTGTTATCCAAAAGGTC 856
Db 577 CAAACTGTTTAGCTCCTATGGCTAAAGCTTTCACGATGCTATTTGTTATCCAAAAGGTC 636
QY 857 TTATGACTACAATCCACGCTTATCTGTTGACCAATGATGCTTACGGACACACCGTG 916
```



```
QY 737 A C A C T A A C C A C G A C A T T C T T G A C G G T A C T G A A A C A G T T A T C T C A G G T G C T T C A T G T A C T A 796
Db 515 A C A C T A A C C A C G A C A T T C T T G A C G G T A C T G A A A C A G T T A T C T C A G G T G C T T C A T G T A C T A 574

QY 797 C A A A C T G T T T A G C T C C T A T G G C T A A A G C T T T C A C G A T G C A T T T G G T A T C C A A A A A G G T C 856
Db 575 C A A A C T G T T T A G C T C C T A T G G C T A A A G C T T T C A C G A T G C A T T C G G T A T T C A A A A A G G T C 634

QY 857 T T A T G A C T A C A A T C C A C G C T T A T A C T G T G A C C A A A T G A T C C T T G A C G G A C C A C A C C G T G 916
Db 635 T T A T G A C T A C A A T C C A C G C T T A C A C T G T G A C C A A A T G A T C C T T G A C G G A C C A C A C C G T G 694

QY 917 G T G G T G A C C T T C G T C G T C G T C G T C G T G C T G C A A A C A A T T G T T C T A A C T C A A C T G G T G 976
Db 695 G T G G T G A C C T T C G T C G T C A C G C G T G T G C T G C A A A C A C A T C G T T C T A A C T C A A C T G G T G 754

QY 977 C T G C T A A A G C T A T C G G T C T T G T T A T C C C A G A A T G A A T G G T A A A C T T G A T G G T G C T G C A C 1036
Db 755 C T G C T A A A G C T A T C G G T C T T G T T A T C C C A G A A C T T A A C G G T A A A C T T G A C G G T G C T G C A C 814

QY 1037 A A C G T G T T C C T G T T C C A A C T G G A T C A G T A A C T G A T T G G T T G T A A C T C T T G A T A A A A A C G 1096
Db 815 A A C G T G T T C C T G T T C C A A C T G G A T C A G T A A C T G A T T G G T T G T A A C T C T T G A C A A A A A C G 874

QY 1097 T T T C T G T T G A C G A A A T C A A C G C T G C T A T G A A A G C T G C T T C A A A C G A C A G T T T C G G T T A C A 1156
Db 875 T T T C T G T T G A C G A A A T C A A C T C T G C T A T G A A A G C T G C T T C A A A C G A T A G C T T C G G T T A C A 934

QY 1157 C T G A A G A T C C A A T T G T T T C T T C A G A T A T C G T A G G C G T G T C A T A C G G T T C A T T G T T T G A C G 1216
Db 935 C T G A A G A T C C A A C T G T T T C T T C A G A T A T C G T A G G C G T A T C A T A C G G T T C A T T G T T T G A C G 994

QY 1217 C A A C T C A A A C T A A A G T T A T G G A A G T T G A C G G A T C A C A A T T G G T T A A A G T T G T A T C A T G G T 1276
Db 995 C A A C T C A A A C T A A A G T A A T G G A A G T T G A C G G A T C A C A A T T G G T T A A A G T T G T A T C A T G G T 1054

QY 1277 A T G A C A A T G A A T G T C T T A C A C T G C T C A A C T T G T T C G T A C A C T T A G T A T T T T G C A A A A A 1336
Db 1055 A T G A C A A C G A A A T G T C T T A C A C T G C T C A A C T T G T A C G T A C T C T T G A G T A C T T C G C A A A A A 1114

QY 1337 T C G C T A A A T A A 1347
Db 1115 T T G C T A A A T A A 1125

RESULT 12
STRPLR
LOCUS STRPLR 1156 bp DNA linear BCT 26-APR-1993
DEFINITION Streptococcus pyogenes plasmin receptor (plr) gene, complete cds.
ACCESSION M95569
VERSION M95569.1 GI:153762
KEYWORDS plasmin binding protein; plasmin receptor.
SOURCE Streptococcus pyogenes
ORGANISM Streptococcus pyogenes
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 1156)
AUTHORS Lottenberg, R., Broder, C.C., Boyle, M.D., Kain, S.J., Schroeder, B.L. and Curtiss, R. III.
TITLE Cloning, sequence analysis, and expression in Escherichia coli of a streptococcal plasmin receptor
JOURNAL J. Bacteriol. 174 (16), 5204-5210 (1992)
MEDLINE 92355491
PUBMED 1322883
COMMENT Original source text: Streptococcus pyogenes DNA.
FEATURES
source Location/Qualifiers
1..1156
/organism="Streptococcus pyogenes"
/mol_type="genomic DNA"
/db_xref="taxon:1314"
48..1125
/gene="plr"

gene
```

```
-35_signal 48..53
/gene="plr"
-10_signal 73..78
/gene="plr"
/note="putative"
RBS 99..105
/gene="plr"
CDS 115..1125
/gene="plr"
/function="plasmin binding protein"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="plasmin receptor"
/protein_id="AAA26953.1"
/db_xref="GI:153763"
/translation="MVVKGINGFGRIGRLAFRRRIQNIIEGVETVRINDLTDPNMLAHL
LKYDTTQGRFDGTVEVKEGGFVNGNFIKVSARDPENIDWATDGVETVLEATGFFAK
KEAAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDILDGTETVIGASCTTNCLAPM
AKALHDAFGIQKGLMTTHAYTDQWILDGPHRGDGLRRARAGAAIVPNSTGAAKAI
GLVIPENGLKLDGAAQRPVPTGVSVELVTLTKVSVDEINSAMKAASNDSTFGYTED
PIVSSDIVGSVSYSLFDATQTKVMEVDGSQLVKVSVWYDNMSYTAQLVRLTLEYFAKI
AK"
ORIGIN
Query Match 58.0%; Score 780.6; DB 1; Length 1156;
Best Local Similarity 94.8%; Pred. No. 2.9e-147;
Matches 807; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 497 TCGATGGTACTGTTGAAGTTAAAGAAGGTGGATTTCGAAGTTAACGGTCAATTTGTTAAAG 556
Db 275 TTGATGGAACAGTTGAAGTTAAAGAAGGTGGATTTCGAAGTTAACGGAAACTTCATCAAAG 334
QY 557 TTTCTGCTGAACGGAAACCAACCAACATTTGACTGGGCTACTGATGGCTAGAAATCGTTC 616
Db 335 TTTCTGCTGAACGGTATCCAGAAACATCGACTGGGCAACTGATGGGTTGAAATCGTTC 394
QY 617 TTGAAGCAACTAGTTTCTTTTGCTAAAAAAGAAGCTGCTGAAAAACACATTACATGCTAACG 676
Db 395 TTGAAGCAACTGGTTTCTTTTGCTAAAAAAGAAGCAGCTGAAAAACACATTACATGCTAACG 454
QY 677 GTGCTAAAAAAGTTGTTATCACAGCTCCTCGTGGAAACGACGCTTAAACAGTTGTTTTC 736
Db 455 GTGCTAAAAAAGTTGTTATCACAGCTCCTCGTGGAAACGATGTTAAACAGTTGTTTTC 514
QY 737 A C A C T A A C C A C G A C A T T C T T G A C G G T A C T G A A A C A G T T A T C T C A G G T G C T T C A T G T A C T A 796
Db 515 A C A C T A A C C A C G A C A T T C T T G A C G G T A C T G A A A C A G T T A T C T C A G G T G C T T C A T G T A C T A 574
QY 797 C A A A C T G T T T A G C T C C T A T G G C T A A A G C T T T C A C G A T G C A T T T G G T A T C C A A A A A G G T C 856
Db 575 C A A A C T G T T T A G C T C C T A T G G C T A A A G C T T T C A C G A T G C A T T T C G G T A T T C A A A A A G G T C 634
QY 857 T T A T G A C T A C A A T C C A C G C T T A T A C T G T G A C C A A A T G A T C C T T G A C G G A C C A C A C C G T G 916
Db 635 T T A T G A C T A C A A T C C A C G C T T A C A C T G T G A C C A A A T G A T C C T T G A C G G A C C A C A C C G T G 694
QY 917 G T G G T G A C C T T C G T C G T C G T G T G T G C T G C A A A C A A C A T T G T T C C T A A C T C A A C T G G T G 976
Db 695 G T G G T G A C C T T C G T C G T G C A C G C G T G T G C T G C A A A C A C A T C G T T C C T A A C T C A A C T G G T G 754
QY 977 C T G C T A A A G C T A T C G G T C T T G T T A T C C C A G A A T T G A A T G G T A A A C T T G A T G G T G C T G C A C 1036
Db 755 C T G C T A A A G C T A T C G G T C T T G T T A T C C C A G A A C T T A A C G G T A A A C T T G A C G G T G C T G C A C 814
QY 1037 A A C G T G T T C C T G T T C C A A C T G G A T C A G T A A C T G A T T G G T T G T A A C T C T T G A T A A A A A C G 1096
Db 815 A A C G T G T T C C T G T T C C A A C T G G A T C A G T A A C T G A T T G G T T G T A A C T C T T G A C A A A A A C G 874
QY 1097 T T T C T G T T G A C G A A A T C A A C G C T G C T A T G A A A G C T G C T T C A A A C G A C A G T T T C G G T T A C A 1156
Db 875 T T T C T G T T C C T G T T C C A A C T G G A T C A G T A A C T G A T T G G T T G T A A C T T G A C A A A A A C G 874
QY 1097 T T T C T G T T G A C G A A A T C A A C G C T G C T A T G A A A G C T G C T T C A A A C G A C A G T T T C G G T T A C A 1156
Db 875 T T T C T G T T G A C G A A A T C A A C T C T G C T A T G A A A G C T G C T T C A A A C G A T A G C T T C G G T T A C A 934
```


SOURCE
ORGANISM Streptococcus pyogenes M1 GAS
Streptococcus pyogenes M1 GAS
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE
AUTHORS 1 (bases 1 to 10053)
Ferretti, J.J., McShan, W.M., Adjic, D., Savic, D., Savic, G., Lyon, K.,
Primeaux, C., Sezate, S.S., Surorov, A.N., Kenton, S., Lai, H., Lin, S.,
Qian, Y., Jia, H.G., Najjar, F.Z., Ren, Q., Zhu, H., Song, L., White, J.,
Yuan, X., Clifton, S.W., Roe, B.A. and McLaughlin, R.E.
Complete genome sequence of an M1 strain of Streptococcus pyogenes
Proc. Natl. Acad. Sci. U.S.A. 98 (8), 4658-4663 (2001)
21192684
MEDLINE
PUBMED 11296296
REFERENCE
AUTHORS 2 (bases 1 to 10053)
Ferretti, J.J., McShan, W.M., Adjic, D., Savic, D., Savic, G., Lyon, K.,
Primeaux, C., Sezate, S.S., Surorov, A.N., Kenton, S., Lai, H., Lin, S.,
Qian, Y., Jia, H.G., Najjar, F.Z., Ren, Q., Zhu, H., Song, L., White, J.,
Yuan, X., Clifton, S.W., Roe, B.A. and McLaughlin, R.E.
Direct Submission
TITLE Submitted (10-APR-2001) Department of Microbiology and Immunology,
JOURNAL University of Oklahoma Health Sciences Center, 940 SL Young Blvd,
Oklahoma City, OK 73104, USA
FEATURES
source
1. .10053
/organism="Streptococcus pyogenes M1 GAS"
/mol_type="genomic DNA"
/strain="SF370"
/serotype="M1"
/db_xref="taxon:160490"
205. .1215
/gene="plr"
/note="synonym: SPY0274"
205. .1215
/gene="plr"
/EC_number="1.2.1.12"
/note="The N-terminal amino acid sequence of this ORF has
been determined from a spot isolated by 2-D gel
electrophoresis from another strain of S. pyogenes. Hogan,
D. A., Whitton, M. M., Rogers, J. and R. A. VanBogelen.
2000. Two-dimensional gel electrophoresis map of
Streptococcus pyogenes proteins. Unpublished data.; Best
Blastp hit = sp|P50467|G3P STRPY GLYCERALDEHYDE
3-PHOSPHATE DEHYDROGENASE (GAPDH) (PLASMINOGEN-BINDING
PROTEIN) (PLASMIN RECEPTOR) >gi|285385|pir||A42963
glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) -
Streptococcus sp >gi|153763|gb|AAA26953.1| (M95569)
plasmin receptor [Streptococcus pyogenes]
/codon_start=1
/transl_table=11
/product="glyceraldehyde-3-phosphate dehydrogenase,
plasmin receptor"
/protein_id="AAK33348.1"
/db_xref="GI:13621550"
/translation="MVKVGVINGFGRIGRLAFRRIQNIEGVETVRINDLTDPMNAHL
LKYDTTQGRFDGTVEKGVNENFIKVSARDPENIDWATDGVETVLEATGFFAK
KEAAEKHLHANGAKKVVITAPGNDVKTVTVENTNHDLDGTETVISGASCTNCLAPM
AKALHDAFGIQGLMTTIHAYTGDMILDGPHRGGLRRARAGAAIVPNSSTCAAKAI
GLVPELNGKLDGAAQRPVPTGVSVELVTLTKNSVDEINAAKKAASNDSPGYTED
PIVSSDIVGSYGSFLFDATQTKVBEVDGSQLVKVSWYDNMSYTAQLVRLTEYPFAKI
AK"
complement(1699. .2439)
/gene="SPY0276"
complement(1699. .2439)
/gene="SPY0276"
/note="Best Blastp hit = pir||H69334 glutamine transport
protein glnQ - Archaeoglobus fulgidus
>gi|2649950|gb|AAB90561.1| (AE001058) glutamine ABC
transporter, ATP-binding protein (glnQ) [Archaeoglobus
fulgidus]"
/codon_start=1
/transl_table=11
/product="putative amino acid ABC transporter (ATP-binding
protein)"
gene
CDS
/protein_id="AAK33349.1"
/db_xref="GI:13621551"
/translation="MSNSIIEIKNLKKSYSNEVLKDISLVNKGVEVISIIGSSGSGK
STLLRSINLLEEPSAQILFHGEDVLAHYNLTHYREKLGVMFQSFNLFENLVLENA
IVAQTTLKRDRAQAEQIAKENLNAGMTEQYWQAKPKQSGQKQKRVAIARALSVNP
EAMLFDEPTSDALDPMEVGEVLKTMQDLAKSGLTMIIVTHEMEFARDVSDRIIFMDKGL
ITEGSPQIFENPTQDRTKEFQRLK"
complement(2432. .4000)
/gene="SPY0277"
complement(2432. .4000)
/gene="SPY0277"
/note="Best Blastp hit = pir||S77250 hypothetical protein
- Synechocystis sp. (strain PCC 6803)
>gi|1652664|dbj|BAA17584.1| (D90907) glutamine-binding
periplasmic protein [Synechocystis sp.]"
/codon_start=1
/transl_table=11
/product="putative glutamine-binding periplasmic protein"
/protein_id="AAK33350.1"
/db_xref="GI:13621552"
/translation="MKLILSLVALALLFGMSRAQANQYLRVGMEEAAYAPFNWTQD
DASGAVPIEGTSQYANGYDVQVAKVAKAMNKELLVKTTSWGLIPALTSKGIDMIA
AGMSPTKERRNEISFSNSSYTSOPVLVVTANGKYADATSLKDFSGAKVTAQQGVVHVN
LLTLKGAKLQTPMGDFSQMRQALTSQVIDAYISERPEAMTAAADSRKMITLKKGF
AVAESDAIAVGMKKDDRMATVQVLEGSQTDRLMALMDMVTQPVKKAEDAKAS
FLGQWMAIFKGNWQFLRGCTGMLLSMVGTTGLFGLLIGFTAPKAKHKAALG
OKLFGWLLTYIEIFRGTPMIVOSMVIYGTQAQAGFISIDRTIAAIFIVSINTGAYMS
EIVRGGIFAVDKGQFKAATALGFTHGQTMRKIVLPQVVRNLPATGNEFVINIKDTSV
LNVISVVELYFSGNTVATQTYQYFQTFTIIAIIYFVLTFTVTRILRYIERFDADTYT
TGANQMQIAEVSNNV"
4198. .6096
/gene="SPY0278"
4198. .6096
/gene="SPY0278"
/note="Best Blastp hit = no blast hits"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK33351.1"
/db_xref="GI:13621553"
/translation="MKKILMTLVLCFSLLGIRIKAADVDYSITNYEGQLLLSKENTAR
FEQKVTYQDTSYNGQYISLGRTHLPAGFAIDQPKVEYVQNGQQVPVSQEFSDLGD
GYRLKLYNAGQAGDKVDKVIWQLHLLLTAYQDVAEIWNWTFISDWDKTLKVSILTVT
PTDIQDNLWAHARGYQYKPKVLEKGNRSYQINAKNVSQGLHAYWDKKGALLGKEPV
DVSTKKNKIIVALETKISRRRLQLLFGKVIPLVEVGFLLWQLIQFTRLKKQFNRYH
LANHTDSYEVPELSPVLVTOAYIQSFAYSPTSASESKLLIPKGVTFEALVQATL
LDLIDQKVLTLTKEEGKAYLEISQDRVTDDEAAFLDMAFGNKVLPVLDQLFSQHYD
ADTIKQLKTKYKGGKLEQEVRSQSQVVKAMKASAAITNNVLETIKKLNLPDTRYQM
TPAEKRSNSVQGLGCLLLILNSGLLIYLAIKESGLALIYALMLVLTMLCLGFYISLKL
DOYKLGLETPEGGVRLHQWQSPKNIIRDIDKFEDVAIEGLVVMNRVLVYATLFGYAK
KVERYLKVHRIALPEVYQAVRPGELSMVMYATPTTFVSSLSAATSSNFSVSSGGGIS
GGGGGGGAF"
6162. .7001
/gene="baca"
/note="synonym: SPY0280"
6162. .7001
/gene="baca"
/note="Best Blastp hit = gb|AAD50462.1|AF169967_5
(AF169967) Baca [Flavobacterium johnsoniae]"
/codon_start=1
/transl_table=11
/product="putative undecaprenol kinase (bacitracin
resistance protein)"
/protein_id="AAK33352.1"
/db_xref="GI:13621554"
/translation="MLIIELLKAIFFGIIIEGITEWLPVSTGTHLILVQEFIRLNQDKA
FIEMENIVIQLGAIIVAVMLIYPERLNPFPQKGTAREVQLTWLWKLKVIACIPSLIIA
VPLDNFEAHFYFMVPIALIVYGIAGFVIEKRNAAQQEFAVTELARMSYKTAFFIGC
FQVLSIVPGTSRSGATILGATILGTSRTVAADTFFLAIPTFMGYSGLKAVKRFLLDGH
HLDFAQVILLVSLTAFVSVLSLAIRFLTDYVKKHDTFTFGKYRIVLGLSLLLIYSFFK
FVF"
7147. .7908
/gene="meca"


```

/organism="Streptococcus pyogenes MGAS315"
/mol_type="genomic DNA"
/strain="MGAS315"
/serotype="M3"
/db_xref="taxon:198466"
/note="group: A"
301..660
/gene="rnpA"
/note="synonym: SpyM3_0175"
301..660
/gene="rnpA"
/note="best blastp match gb|AAK33323.1| (AE006491)
putative ribonuclease P protein component [Streptococcus
pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative ribonuclease P protein component"
/protein_id="AAM78782.1"
/db_xref="GI:21903895"
/translation="MKRTYRVKREKDFQALFKDGKSTANRKFVYHLNRGQDHFVRVGI
SVGKIGNAVTRNAVTRKIRHVIMALGHQLKSEDFVVIARKGVESLEYQELQQNLHHV.
LKLAQLLEKGFSEBKH"
626..1453
/gene="SpyM3_0176"
626..1453
/gene="SpyM3_0176"
/note="best non-GAS blastp hit: gb|AAK76106.1| (AE007493)
SpoIIJ family protein [Streptococcus pneumoniae TIGR4]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAM78783.1"
/db_xref="GI:21903896"
/translation="MRKVLVRKKNKIARIIVPLVLLLVACGRGEVTAQSSSGWDQLVY
LFARAIQWLSFPGSIGVGIIILFTLTIRMLMPLFNMQIKSSQKQDIQPELRELQTKY
AGKDTQTRMKLAHESQALYKKGYNPYASLLPLLQMPVMIALFQALTRVSLKGTGF
LWVELAQHDHLXILPVAAVFTFLSTWLTNLAKEKNVMTVMIVVPLMIFFMGFNL
ASGVLYTWVSNAFQVVQVLLLLNNPFKIIAERQRLANEERRLRARRKAMKRRK"
1465..2379
/gene="SpyM3_0177"
1465..2379
/gene="SpyM3_0177"
/note="best non-GAS blastp hit: gb|AAK76105.1| (AE007493)
jag protein, putative [Streptococcus pneumoniae TIGR4],
and gb|AAL00654.1| (AE008549) hypothetical protein
[Streptococcus pneumoniae R6]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAM78784.1"
/db_xref="GI:21903897"
/translation="MVLFTGKTVEAETGLQELGLPRLKAHIKVISKEKKGFLGFGK
KPAQVDIEGSDPTQVYKADKKATRGVPEDINRQNAFVNSADVEPEEIKATORLEAED
TKVPLMSDQAIIPSNLAETVTETKAQPSIPVEESEVPQDAGNDGFSKDIEKAAQ
EVSDYVTKIIEYMDIEATVETSNRRQINLIQETPEAGRVIGYHGKVLKSLQLLAQNF
LHDYRSKNFSVSLNVHDYVHRFTETLIDFTQKVAKRVLSEGGDYTMDDPMSNSERKIVH
KTVSSIEGVDSYSEGNDPNRYVVVSLQR"
2694..2828
/gene="rpmH"
/note="synonym: SpyM3_0178"
2694..2828
/gene="rpmH"
/note="best blastp hit: gb|AAK33326.1| (AE006491) 50S
ribosomal protein L34 [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L34"
/protein_id="AAM78785.1"
/db_xref="GI:21903898"
/translation="MKRTYQPSKIRRQKKGFRHRMSTKNGRRVLAARRRKGKRVLSA
"
3102..3806
/gene="SpyM3_0179"

```

```

3102. .3806
/gene="SpyM3_0179"
/note="best blastp hit: gb|AAK33327.1| (AE006492) putative
N-acetylmannosamine-6-P epimerase [Streptococcus pyogenes
M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative N-acetylmannosamine-6-P"
/protein_id="AAM78786.1"
/db_xref="GI:21903899"
/translation="MPDKPTKELMEQLKGGIIVSCQALPGEPLYSETGGIMPLLAKE
AQEAGVGIGIRANSVRDIDKEIQAITDPIIGIIKDYPPQEPFITATMAEVDQLAALNI
AVIAMDCKDRHDGLDIASFIRQVKEKYPNQLLMADISTFDEGLVAHQAGIDFVGTT
LSGYTPYGRQEAGDPDVALIEALCKAGIAVIAEGKIHSPEEAKKINDLIGVAGIIVGGAI
TRPKEIAERFIEALKS"
3855. .4913
/gene="SpyM3_0180"
3855. .4913
/gene="SpyM3_0180"
/note="best blastp hit: gb|AAK33328.1| (AE006492) putative
sugar transporter sugar binding lipoprotein [Streptococcus
pyogenes M1 GAS]; possible pseudogene aligns with
N-terminal portion (amino acids 1-352)"
/codon_start=1
/transl_table=11
/product="putative sugar transporter sugar binding
lipoprotein"
/protein_id="AAM78787.1"
/db_xref="GI:21903900"
/translation="NMKKLASLVMLGASVLGLAACGGKSQKEAGASKSDTAKTEITW
WAFPVFTEKAEDGVGYEKKLIAAFKANPEIKVLETFIDTSGPEKITTAEAGTA
PDVLFDPAGRIIOYKNGKGLADLNDLFTKEFTKDVNDKLIQASKAGDTAYMYPISSA
PFYMALNKMKFKDAGVLDLVKEGWTDDFEKVLKALKDKGYNPGSFFANGQGGDQGR
AFFANLYSSHITDDKVTKYTTDDANSIKAMTKISNWKDGLMMNGSQYDGSADIQNFPA
NGQTSFTILWAPAQPGIOAKLLEASKVDYLEIPFPSDDGKPELEYLVNGFAVFNKDE
QKVAASKTFIQFIADDK"
4989. .5174
/gene="SpyM3_0181"
4989. .5174
/gene="SpyM3_0181"
/note="best blastp hit: gb|AAK33328.1| (AE006492) putative
sugar transporter sugar binding lipoprotein [Streptococcus
pyogenes M1 GAS]; possible pseudogene aligns with
C-terminal portion (amino acids 379-439)"
/codon_start=1
/transl_table=11
/product="putative sugar transporter sugar binding
lipoprotein"
/protein_id="AAM78788.1"
/db_xref="GI:21903901"
/translation="MEKIAEWTKFSPYNTIDGFAEMRTLWFPVMQVAINSGDEKPED
ALKAPTEKANKTIKKTQ"
5299. .6165
/gene="SpyM3_0182"
5299. .6165
/gene="SpyM3_0182"
/note="best blastp hit: gb|AAK33329.1| (AE006492) sugar
transport system (permease) [Streptococcus pyogenes M1
GAS]"
/codon_start=1
/transl_table=11
/product="putative sugar transport system (permease)"
/protein_id="AAM78789.1"
/db_xref="GI:21903902"
/translation="MRETLISYAF LAPVLVFFVIFVLIPMIMGFVTSFFNYSMTDFTF
VGFANYARMFQDPIFKSLMNTLIIIGSVPVVVFSLFVAAKTYDKNVVARSFYRAV
FPLPVVTGSVAIVYVNMKIYDPMSGILNYLVKAVHIEQNI SWLGDKHWALLAI IAIL
LITSVQGP IILYIAAMNIDNSLMEARVDGATEFQVFNWKPSLLPTLYIAVITT
INSFQCFALIQLLTSGGNYSTSTLMYYLYEKAFLSEGYANTMGVFLAVMIAIISF
AQFKILGNDVEY"
6178. .7008
/gene="SpyM3_0183"
6178. .7008

```

```
/gene="SpyM3_0183"  
/note="best blastp hit: gb|AAK33330.1| (AE006492) probable  
sugar transport protein (permease) [Streptococcus pyogenes  
M1 GAS]"  
/codon_start=1  
/transl_table=11  
/product="putative sugar transport protein (permease)"  
/protein_id="AAM78790.1"  
/db_xref="GI:21903903"  
/translation="MKKKLTASDVLTVTILCILTFIFPFYWIMTGAFKAQADTIM  
IPQWPKAPTIFENFKALVIONPALKWLWNSVFISVTFMLVCGTSSLAGYALAKRRF  
YQRIILFSIFIAAVALPKQVVLVPLVIRVFMGHDHDLAAILPLVGVFPFVFLMKQF  
SENIPTLELESADKIDGCEIRTFENVAFFPIVKPGEAALAIPTFINTWYDFMLVMLT  
SRENLTISLGVATMQAEMATNYGLIMAGAAVAAPVITVFLVFKSFTQGITMGAVKG  
"  
7165..7827  
/gene="SpyM3_0184"  
7165..7827  
/gene="SpyM3_0184"  
/note="best non-GAS blastp hit: gb|AAK75756.1| (AE007460)  
hypothetical protein [Streptococcus pneumoniae TIGR4]"  
/codon_start=1  
/transl_table=11  
/product="conserved hypothetical protein"
```

Query Match		58.0%;	Score 780.6;	DB 1;	Length 50029;
Best Local Similarity		94.8%;	Pred. No. 2.4e-147;		
Matches 807;		Conservative 0;	Mismatches 44;	Indels 0;	Gaps 0;
QY	497	TCGATGGTACTGTTGAAGTTAAAGAAGGTGGATTGAAAGTTAACGGTCAATTTGTTAAAG	556		
Db	25967	TTGATGGAACAGTTGAAGTTAAAGAAGGTGGATTGAAAGTTAAACGGAAACTTCATCAAAG	26026		
QY	557	TTTCTGCTGAACCGCAACAGCAACACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC	616		
Db	26027	TTTCTGCTGAACGCTGATCCAGAAACACATCGACTGGCAACTGATGGGTTGAAATCGTTC	26086		
QY	617	TTGAAGCAACTAGTTCTTTGCTAAATAAGAGAGCTGCTGAAAAACACATTACATGCTAACG	676		
Db	26087	TTGAAGCAACTGGTTCTTTGCTAAATAAGAGAGAGCTGAAAAACACATTACATGCTAACG	26146		
QY	677	GTGCTAAAAAAGTTGTTATCACAGCTCCTGTTGGAACGACGTTTAAACAGTTGTTTCA	736		
Db	26147	GTGCTAAAAAAGTTGTTATCACAGCTCCTGTTGGAACGATGTTTAAACAGTTGTTTCA	26206		
QY	737	ACACTAACCCAGCAGATTCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA	796		
Db	26207	ACACTAACCCAGCAGATTCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA	26266		
QY	797	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTACAGATGCAATTTGGTATCCAAAAAGGTC	856		
Db	26267	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTACAGATGCAATTTGGTATCCAAAAAGGTC	26326		
QY	857	TTATGACTACAAATCCACGGCTTATACCTGGTGACCAAAATGATCCTTGACGGACCAACCCGTG	916		
Db	26327	TTATGACTACAAATCCACGGCTTATACCTGGTGACCAAAATGATCCTTGACGGACCAACCCGTG	26386		
QY	917	GTGGTGACCTTCGTGCTGCTCGTGTGCTGCTGCAAAACATTTGTTCCCTAACTCAACTGGTG	976		
Db	26387	GTGGTGACCTTCGTGCTGCTCACGCGCTGGTGTGCTGCAAAATATCGTTCCCTAACTCAACTGGTG	26446		
QY	977	CTGCTAAAGCTATCGGTCTTGTATCCAGAAATGAATGGTAAACTTGATGGTGTGCAC	1036		
Db	26447	CTGCTAAAGCTATCGGTCTTGTATCCAGAAATGAATGGTAAACTTGATGGTGTGCAC	26506		
QY	1037	AACGTGTTCCCTGTTCCAACTGGATCAGTAAGTGGTGGTGTGTAACCTTGATAAAAAACG	1096		
Db	26507	AACGTGTTCCCTGTTCCAACTGGATCAGTAAGTGGTGGTGTGTAACCTTGACAAAAACG	26566		
QY	1097	TTTCTGTTGACGAATCAACGCTGCTATGAAGCTGCTTCAAACGACAGTTTCGGTTACA	1156		
Db	26567	TTTCTGTTGACGAATCAACGCTGCTATGAAGCTGCTTCAAACGATAGCTTCGGTTACA	26626		

QY	1157	CTGAAGATCCAATTTCTTTCAGATATCGTAGGCGTGTATACCGTTTCATTGTTGACG	1216
Db	26627	CTGAAGATCCAATCGTTTCTTTCAGATATCGTAGGCGTATATACCGTTTCATTGTTGACG	26686
QY	1217	CAACTCAAACTAAAGTTATGGAAGTTGACGGATCAACAATTGGTTAAAGTTGTATCATGGT	1276
Db	26687	CAACTCAAACTAAAGTAATGGAAGTTGACGGATCAACAATTGGTTAAAGTTGTATCATGGT	26746
QY	1277	ATGACAATGAAATGCTTACACTGCTCAACTTTGTTGCTACACTTGAGTATTTTGCAAAAA	1336
Db	26747	ATGACAACGAAATGCTTACACTGCTCAACTTTGTTGCTACACTCTTGAGTACTCTTGAGTACTTCGCAAAAA	26806
QY	1337	TCGCTAAATAA	1347
Db	26807	TTGCTAAATAA	26817

Search completed: March 31, 2004, 00:58:07
Job time : 5274 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 18:05:29 ; Search time 566 Seconds
(without alignments)
10110.111 Million cell updates/sec

Title: US-10-650-369-21
Perfect score: 1347
Sequence: 1 atgaaaaaaataacagggat.....ttgcaaaaatcgctaaataa 1347

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212499041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1347	100.0	1347	6	ABA91327 Streptoco
2	815.8	60.6	1011	6	ABA91328 Streptoco
3	815.8	60.6	1011	6	ABA91248 Streptoco
4	780.6	58.0	1011	7	ACA50325 Prokaryot.
5	780.6	58.0	1125	2	AAQ70705 Plasmin r
6	780.6	58.0	1125	4	AAC66140 Streptoco
7	777.6	57.7	1008	6	ABN70591 Streptoco
8	777.6	57.7	1035	6	ABN69737 Streptoco
9	702.2	52.1	1011	6	ABA91329 Streptoco
10	702.2	52.1	1011	6	ABA91249 Streptoco
11	699	51.9	110000	6	ABN71527_17 Continuation (18 o
12	696	51.7	1008	6	ABN71389 Streptoco
13	683	50.7	1080	4	AAS55858 Streptoco
14	678.2	50.3	1011	6	ABA91332 Streptoco
15	678.2	50.3	1011	6	ABA91252 Streptoco
16	678.2	50.3	1080	4	AAS55434 Streptoco
17	678.2	50.3	1080	7	ACA50132 Prokaryot
18	678.2	50.3	7571	2	AAV52135 Streptoco
19	678.2	50.3	110000	7	ABS56454_18 Continuation (19 o
20	676.6	50.2	1011	6	ABA91330 Streptoco
21	676.6	50.2	1011	6	ABA91250 Streptoco
22	675.2	50.1	1005	7	ABX07806 S. pneumo
23	660.2	49.0	1000	2	AAV27350 Streptoco

24	660.2	49.0	1000	6	ABQ84818 S. pneumo
25	660.2	49.0	1000	9	ADC45134 S. pneumo
26	640.8	47.6	5448	2	AAV65229 DNA encod
27	639.8	47.5	1011	6	ABA91331 Streptoco
28	639.8	47.5	1011	6	ABA91251 Streptoco
29	620.6	46.1	1011	7	ACA48200 Prokaryot
30	567.8	42.2	65589	6	ABA90521_23 Continuation (24 o
31	495.6	36.8	1002	9	ADC91488 E. faeciu
32	481.4	35.7	110000	6	ABA90521_05 Continuation (6 of
33	479.6	35.6	1002	4	AAS53113 Enterococ
34	478	35.5	999	7	AAS53113 Enterococ
35	478	35.5	4023	2	ACA33240 Prokaryot
36	478	35.5	4023	6	AAX13023 Enterococ
37	454	33.7	110000	6	ABS98818 Enterococ
38	454	33.7	110000	6	ABQ67196_1 Continuation (2 of
39	453.6	33.7	1005	7	ACA27465_25 Continuation (26 o
40	450.8	33.5	1011	7	ACA36438 Prokaryot
41	450.8	33.5	110000	6	ABA3041_25 Continuation (26 o
42	447.6	33.2	4230	6	ABQ70965 Listeria
43	426.8	31.7	1011	4	AAH53908 S. epider
44	426.8	31.7	1047	6	ABN93213 Staphyloc
45	426.8	31.7	4642	4	AAH54726 S. epider

ALIGNMENTS

RESULT 1
ABA91327
ID ABA91327 standard; DNA; 1347 BP.
XX
AC ABA91327;
XX
DT 29-AUG-2003 (revised)
DT 08-APR-2002 (first entry)
XX
DE Streptococcus GapC multiepitope fusion PolyGap4 DNA.
XX
KW PolyGap4; GapC; plasmin binding protein; epitope; infection; vaccine;
KW immunisation; mastitis; therapy; gene; ds.
XX
OS Streptococcus dysgalactiae.
OS Streptococcus agalactiae.
OS Streptococcus parauberis.
OS Chimeric.
XX
PN WO200196379-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000836.
XX
PR 12-JUN-2000; 2000US-0211247P.
XX
PA (UJSA-) UNIV SASKATCHEWAN.
XX
PI Potter AA, Perez-Casal J, Fontaine M;
XX
DR WPI; 2002-098051/13.
DR P-PSDB; AAM50664.
XX
PT Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
PT parauberis, or S. iniae GapC protein useful for treating mastitis in
PT vertebrates.
XX
PS Example 4; Fig 6A-C; 116pp; English.
XX
CC The present sequence is that of a chimeric gene encoding a multiple
CC epitope fusion protein (see AAM50664), designated PolyGap4, comprising
CC the entire amino acid sequence of the Streptococcus dysgalactiae GapC
CC plasmin binding protein in addition to unique amino acid sequences from
CC the Streptococcus parauberis and Streptococcus agalactiae GapC proteins.

The gapC gene sequences used to construct the chimeric gene were prepared by PCR amplification of selected polynucleotide sequences from the genomic gapC genes using the primers given in ABA91335-42. After assembly, the chimeric gene was inserted into plasmid pAA555. PolyGal4 is an example of novel GapC multiple epitope fusion proteins of the invention that comprise epitopes from 1 or more of *S. dysgalactiae*, *S. agalactiae*, *S. parauberis*, *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-69). The multiple epitope fusion proteins are produced using claimed host-vector systems and used in claimed vaccines for treating or preventing a bacterial infection in a vertebrate, especially a streptococcal infection, and particularly mastitis. They are also used in claimed methods of detecting *Streptococcus* antibodies. The multiple epitope protein is capable of eliciting broad immunity against a variety of streptococcal infections while minimising the number of antigens present in the final formulation and concomitantly reducing production costs. (Updated on 29-AUG-2003 to standardise OS field)

Sequence 1347 BP; 401 A; 251 C; 297 G; 398 T; 0 U; 0 Other;

Query Match 100.0%; Score 1347; DB 6; Length 13

Query Match 100.0%; Score 1347; DB 6; Length 1347;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1347; Conservative 0; Mismatches 0; Indels 0;

QY	1	ATGAAAAAATAACAGGATATTTATTTGCTTCTTGCAGTCATTATTCGTCTGCATGC	60
DB	1	ATGAAAAAATAACAGGATATTTATTTGCTTCTTGCAGTCATTATTCGTCTGCATGC	60
QY	61	CAGGCAAACTACGGATCCGGTATGGTAGTTAAAGTTGGTATTAAACGGTTTCGGTCGTATC	120
DB	61	CAGGCAAACTACGGATCCGGTATGGTAGTTAAAGTTGGTATTAAACGGTTTCGGTCGTATC	120
QY	121	GGACGTCCTTGCAATCCGTCGTATTCAAAATGTTGAAGGTGTTGAAGTAACTCGTATCAAC	180
DB	121	GGACGTCCTTGCAATCCGTCGTATTCAAAATGTTGAAGGTGTTGAAGTAACTCGTATCAAC	180
QY	181	GACCTTACAGATCCAAACATGCTTGCAACACTGTTGAAATACGATACAACTCAAGGACGT	240
DB	181	GACCTTACAGATCCAAACATGCTTGCAACACTGTTGAAATACGATACAACTCAAGGACGT	240
QY	241	TTTGACGGAACTGTTGAAGTTAAAGAGGTGGATTTGAAGTAAACGGAAACTTCATCAAA	300
DB	241	TTTGACGGAACTGTTGAAGTTAAAGAGGTGGATTTGAAGTAAACGGAAACTTCATCAAA	300
QY	301	GTTTCTGCTGAACGTGATCCAGAAAAACATCGACTGGGCAACTGACGGTGTGAAATCGTT	360
DB	301	GTTTCTGCTGAACGTGATCCAGAAAAACATCGACTGGGCAACTGACGGTGTGAAATCGTT	360
QY	361	CTGGAAGCACTCGAGGGTACTGTAGAGTTAAAGATGGTGGATTTGACGTTAACGGAAAA	420
DB	361	CTGGAAGCACTCGAGGGTACTGTAGAGTTAAAGATGGTGGATTTGACGTTAACGGAAAA	420
QY	421	TTCAATTAAAGTTTCTGCTGAAAAAGATCCAGAACAAATTGACTGGGCAACTGACGGTGT	480
DB	421	TTCAATTAAAGTTTCTGCTGAAAAAGATCCAGAACAAATTGACTGGGCAACTGACGGTGT	480
QY	481	GAAATCGTTCCTTGAAATCGATGGTACTGTTGAAGTTAAAGAAAGTGGATTTCGAAGTTAAC	540
DB	481	GAAATCGTTCCTTGAAATCGATGGTACTGTTGAAGTTAAAGAAAGTGGATTTCGAAGTTAAC	540
QY	541	GGTCAATTGTTAAAGTTTCTGCTGAACGCGAACCCAGCAACATTGACTGSSCTACTGAT	600
DB	541	GGTCAATTGTTAAAGTTTCTGCTGAACGCGAACCCAGCAACATTGACTGSSCTACTGAT	600
QY	601	GGCGTAGAAATCGTTCCTTGGAAGCAACTAGTTTCTTTGCTAAAAAAGAGCTGCTGAAAAA	660
DB	601	GGCGTAGAAATCGTTCCTTGGAAGCAACTAGTTTCTTTGCTAAAAAAGAGCTGCTGAAAAA	660
QY	661	CACCTACATGCTAACCGTGCTAAAAAGTTGTTATCACAGCTCCTGGTGGAAAAACGACGTT	720
DB	661	CACCTACATGCTAACCGTGCTAAAAAGTTGTTATCACAGCTCCTGGTGGAAAAACGACGTT	720
QY	721	AAAACAGTTGTTTTCAACACTAACCCACGACATTCCTTGACGGTACTGAAACAGTTATCTCA	780

RESULT 2
ABA91328

ABA91328
ID ABA91328 standard: DNA: 1011 bp.

AC ABA91328;

XX
DT 08-APR-2002 (first entry)

XX DE STREPTOCOCCUS DYSAGALACTIAE

XXIX

immunisation; mastitis; therapy; gene; ds.

05 Streptococcus dysgalactiae.

XX PN WO200196379-A2.

XX
 20 DEC 2001[illegible]































1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PA (UySA-) UNIV SASKATCHEWAN.
XX

PI Potter AA, Perez-Casal J, Fontaine M,

DR WPI; 2002-098051/13.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

RESULT 4

ACA50325
ID ACA50325 standard; DNA: 1011 BP.

XX
AC ACA50325;

19-JUN-2003 (first entry)

XXII

DE	Prokaryotic essential gene #31982.	
XX		
KW	Antisense; ds; prokaryotic essential gene; cell proliferation;	
KW	drug design; gene.	
XX		
OS	Streptococcus pyogenes.	
XX		
PN	WO200277183-A2.	
XX		
PD	03-OCT-2002.	
XX		
PF	21-MAR-2002; 2002WO-US009107.	
XX		
PR	21-MAR-2001; 2001US-00815242.	
PR	06-SEP-2001; 2001US-00948993.	
PR	25-OCT-2001; 2001US-0342923P.	
PR	08-FEB-2002; 2002US-00072851.	
PR	06-MAR-2002; 2002US-0362699P.	
XX		
PA	(ELIT-) ELITRA PHARM INC.	
XX		
PI	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;	
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;	
XX		
DR	WPI; 2003-029926/02.	
DR	P-PSDB; ABU46455.	
XX		
PT	New antisense nucleic acids, useful for identifying proteins or screening	
PT	for homologous nucleic acids required for cellular proliferation to	
PT	isolate candidate molecules for rational drug discovery programs.	
XX		
PS	Claim 14; SEQ ID NO 38195; 1766pp; English.	

Sequence 1011 BP: 300 A: 206 C: 213 G: 292 T: 0 U: 0 Other: 0

Query Match 58.0%; Score 780.6; DB 7; Length 1011;
Best Local Similarity 94.8%; Pred. No. 8.5e-198;

QY 497 TCGATGGTACTGTTGAAGTTAAAGAAGGTGGATT CGAAGTTAACCGGTCAATTGTTAAAG 556

Db	161	TTGATGGAAACAGTTGAAGTTAAAGAAAGGTGGATTGAAGTAAACGGAAACTTCATCAAAG	220
QY	557	TTTCTGCTGAACGGAAACGAGCAAAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTTC	616
Db	221	TTTCTGCTGAACGTTGATCCAGAAAAACATCGACTGGGCAACTGATGGGTTGAAATCGTTTC	280
QY	617	TTGAAGCAAACTAGTTTCTTTTGCTAAAAAAGAAAGCTGCTGAAAAACACATTACATGCTAAACG	676
Db	281	TTGAAGCAAACTGGTTTCTTTTGCTAAAAAAGAAAGCAAGCTGAAAAACACATTACATGCTAAACG	340
QY	677	GTGCTAAAAAAGTTGTTATCACAGCTCCTGGTGGAAAAACGAGTTAAAAACAGTTGTTTCA	736
Db	341	GTGCTAAAAAAGTTGTTATCACAGCTCCTGGTGGAAAAACGAGTTAAAAACAGTTGTTTCA	400
QY	737	ACACTAACCCACGACATCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA	796
Db	401	ACACTAACCCACGACATCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA	460
QY	797	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTACCGATGCATTTGGTATCCAAAAAGGTC	856
Db	461	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTACCGATGCATTTGGTATCCAAAAAGGTC	520
QY	857	TTATGACTACAATCCACGCTTATACCTGGTGACCAAAATGATCCTTGACGGACACACCGTG	916
Db	521	TTATGACTACAATCCACGCTTATACCTGGTGACCAAAATGATCCTTGACGGACACACCGTG	580
QY	917	GTGGTGACCTTCGTCGTGCTCGTGTGCTGGTGCTGCAAAACATTGTTCTTAACCTCAACTGGTG	976
Db	581	GTGGTGACCTTCGTCGTGCTCGTGTGCTGGTGCTGCAAAATATCGTTCTTAACCTCAACTGGTG	640
QY	977	CTGCTAAAGCTATCGGTCTTGTTATCCAGAAATGAATGGTAAACCTTGATGGTCTGCAC	1036
Db	641	CTGCTAAAGCTATCGGTCTTGTTATCCAGAACTTAACGGTAAACCTTGACGGTCTGCAC	700
QY	1037	AACGTGTTCCCTGTTCCAACTGGATCAGTAACCTGAGTTGGTTGTAACCTCTTGATAAAAAACG	1096
Db	701	AACGTGTTCCCTGTTCCAACTGGATCAGTAACCTGAGTTGGTTGTAACCTCTTGACAAAAACG	760
QY	1097	TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAAACGACAGTTTCGGTTACA	1156
Db	761	TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAAACGATAGCTTCGGTTACA	820
QY	1157	CTGAAGATCCAATTGTTTCTTCAGATATCGTAGGCGTGCATACGGTTCAATTGTTTGACG	1216
Db	821	CTGAAGATCCAATCGTTTCTTCAGATATCGTAGGCGTGCATACGGTTCAATTGTTTGACG	880
QY	1217	CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTGGTTAAAGTTGTATCATGGT	1276
Db	881	CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTGGTTAAAGTTGTATCATGGT	940
QY	1277	ATGACAAATGAAATGCTTTACACTGCTCAACTTGTTCGTACACTTGAGTATTTGCAAAA	1336
Db	941	ATGACAAAGAAATGCTTTACACTGCTCAACTTGTTCGTACACTTGAGTATTTGCAAAA	1000
QY	1337	TCGCTAAATAA	1347
Db	1001	TTGCTAAATAA	1011

RESULT 5
AAO70705

AAQ70705 standard; DNA; 1125 BP.

XX
AC
AAQ70705;

DT 25-MAR-2003 (revised)

DT 22-FEB-1995 (first entry)

DE Plasmin receptor DNA.

XX Plasmin receptor; isolate 64/14; plasmin; alpha-2-antiplasmin;
KW plasminogen activator; bleeding; reocclusion; thrombosis;
KW pulmonary embolism; clots; ss.

XX	Streptococcus pyogenes.
OS	
XX	
FH	Key
FT	Location/Qualifiers
FT	115. .1125
FT	/*tag= a
FT	/product= "plasmin receptor"

US5328996-A.

12-JUL-1994.

10-AUG-1992; 92US-00928462.

29-MAR-1989; 89US-00330849.

16-MAY-1990; 90US-00524411.
(UYFL) UNIV FLORIDA RES FOUND INC.

Von Mering G, Broder C, Boyle MDP, Lottenberg R;
WPI; 1994-225327/27.
P-PSDB; AAR56486.

New DNA encoding bacterial plasmin receptor - useful as thrombolytic agents, used with plasminogen activator or bound to plasmin, also useful in vaccines against bacterial infection.

Claim 2; Col 27-30; 19pp; English.

Query.Match 58.0%; Score 780.6; DB 2; Length 1125;

Best Local Similarity 94.8%; Pred. No. 8.9e-198;

Matches 807; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 497 TCGATGGTACTGTTGAAGTTAAAGAAGGTGGATTCTGAAGTTACGGTCAATTGTGTTAAAG 556

557 TTTCTGCTGAACGGGAACCAAGCAAA CATTGACTGGGCTACTGATGGCGTAGAAATCGTTC 616

335 TTTCTGCTGAACGTGATCCAGAAAAACATCGACTGGGCAACTGATGGGTTGAAATCGTTC 394

617 TTGAAGCAACTAGTTCTTTGCTAAAAAGAAAGCTGCTGTAACCTTACATGCTAACG 676

395 TTGAAGCAACTGGTTCTTTGCTAAAGAGCAGCTGAAAAACACTTACATGCTAACG 454

677 GTGGCTA A A A A A A G T T G T T A T C A C A G C T C C T G T G G A A A C G A C G T T A A A A C A G T T G T T T C A 736

DP 155 CTCCTATAAAGCTTCTTTATCAACCGCTCCTCGTGGAAACGATGTTAAACAGTTGTTTCA 514

732 AGAGCTAGCCAGCATTCTTCACCGGTATCTCAACAGTTATCTCAGGTGCTTCATGTA CTA 796

[illegible]

CAACGCGCCGCCGCCTCAGTGGTTATCAAAGAAGTC856

634

Db 995 CAACTCAAACTAAAGTAATGGAAGTTGACGGATCACAAATGGTTAAAGTTGTATCATGGT 1054
QY 1277 ATGACAATGAAATGCTTACACTGCTCAACTTGTTCGTACACTTGAGTATTTTGCAAAAA 1336
Db 1055 ATGACAACGAAATGCTTACACTGCTCAACTTGTACGTACTCTTGAGTACTTCGCAAAAA 1114
QY 1337 TCGCTAAATAA 1347
Db 1115 TTGCTAAATAA 1125
RESULT 7
ID ABN70591 standard; DNA; 1008 BP.
XX
AC ABN70591;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 9095.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
DR P-PSDB; ABP29960.
XX
PT New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
PS Claim 7; Page 4030; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 1008 BP; 298 A; 206 C; 213 G; 291 T; 0 U; 0 Other;
Query Match 57.7%; Score 777.6; DB 6; Length 1008;

Best Local Similarity 94.8%; Pred. No. 5.4e-197;
Matches 804; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 497 TCGATGTTACTGTTGAAGTTAAAGAGGTGGATTTCGAAGTTAAACGGTCAATTTGTTAAAG 556
Db 161 TTGATGGAACAGTTGAAGTTAAAGAGGTGGATTTCGAAGTTAAACGGGAACTTCATCAAAG 220
QY 557 TTTCTGCTGAACGGCAACAGCAACANTGACTGGGCTACTGATGGCGTAGAATCGTTTC 616
Db 221 TTTCTGCTGAACGGTGTATCCAGAAAACATCGACTGGGCAACTGATGGGGTTGAATCGTTTC 280
QY 617 TTGAAGCAACTAGTTTCTTTTGCTAAAAAAGAAAGCTGCTGAAAAACACTTACATGCTAACG 676
Db 281 TTGAAGCAACTAGTTTCTTTTGCTAAAAAAGAAAGCAGCTGAAAAACACTTACATGCTAACG 340
QY 677 GTGCTAAAAAAGTTGTTATCACAGCTCCTGTTGGAACAGCTTAAAAACAGTTGTTTCA 736
Db 341 GTGCTAAAAAAGTTGTTATCACAGCTCCTGTTGGAACAGCTTAAAAACAGTTGTTTCA 400
QY 737 AACTAACACACAGCATTTCTTGACGGTACTGAAACAGTTTATCTCAGGTGCTTCACTGACTA 796
Db 401 AACTAACACACAGCATTTCTTGACGGTACTGAAACAGTTTATCTCAGGTGCTTCACTGACTA 460
QY 797 CAAACTGTTTAGCTCCTATGCTTAAAGCTCTTTCACGATGCTATTTGTTATCCAAAAAGGTC 856
Db 461 CAAACTGTTTAGCTCCTATGCTTAAAGCTCTTTCACGATGCTATTTGTTATCCAAAAAGGTC 520
QY 857 TTATGACTACAATCCACGCTTATCTGTTGACCAATGATCTTACGGACCAACCGTG 916
Db 521 TTATGACTACAATCCACGCTTATCTGTTGACCAATGATCTTACGGACCAACCGTG 580
QY 917 GTGCTGACCTTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 976
Db 581 GTGCTGACCTTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 640
QY 977 CTGCTAAAGCTATCGGTCCTGTTATCCAGAAATGATGTTGAATGTTGTTGTTGTCGTCAC 1036
Db 641 CTGCTAAAGCTATCGGTCCTGTTATCCAGAAATGATGTTGAATGTTGTTGTCGTCAC 700
QY 1037 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACCTTTGATAAAAAACG 1096
Db 701 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACCTTTGATAAAAAACG 760
QY 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGTTTACA 1156
Db 761 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGTTTACA 820
QY 1157 CTGAAGATCCAATTTGTTCTTTCAGATATCGTAGGCTGTATACGTTTCACTGTTGACG 1216
Db 821 CTGAAGATCCAATCGTTTCTTTCAGATATCGTAGGCTGTATACGTTTCACTGTTGACG 880
QY 1217 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCAAAATGGTTAAAGTTGATCATGGT 1276
Db 881 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCAAAATGGTTAAAGTTGATCATGGT 940
QY 1277 ATGACAATGAAATGTTTACACTGCTCAACTTGTTCGTACACTTGTAGTATTTTGCAAAAA 1336
Db 941 ATGACAACGAAATGTTTACACTGCTCAACTTGTTCGTACACTTGTAGTATTTTGCAAAAA 1000
QY 1337 TCGCTAAA 1344
Db 1001 TTGCTAAA 1008
RESULT 8
ABN69737
ID ABN69737 standard; DNA; 1035 BP.
XX
AC ABN69737;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 7387.

XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
PT parauberis, or S. iniae GapC protein useful for treating mastitis in
PT vertebrates.
XX
PS Example 3; Fig 2A-B; 116pp; English.
XX
CC The present sequence is that of the coding region of the gene encoding
CC the GapC plasmin binding protein, AgalGapC (see AAM50666), of
CC Streptococcus agalactiae ATCC 27541. The gene was obtained by PCR
CC amplification of chromosomal DNA using the primers given in ABA91333-34.
CC The invention relates to novel GapC multiple epitope fusion proteins that
CC comprise epitopes from 1 or more of Streptococcus dysgalactiae,
CC Streptococcus agalactiae, Streptococcus parauberis, Streptococcus uberis
CC and Streptococcus iniae (see AAM50665-69). A claimed example is PolyGap4
CC (see AAM50664). The multiple epitope fusion proteins are produced using
CC claimed host-vector systems and used in claimed vaccines for treating or
CC preventing a bacterial infection in a vertebrate, especially a
CC streptococcal infection, and particularly mastitis. They are also used in
CC claimed methods of detecting Streptococcus antibodies. The multiple
CC epitope proteins are capable of eliciting broad immunity against a
CC variety of streptococcal infections while minimising the number of
CC antigens present in the final formulation and concomitantly reducing
CC production costs
XX
SQ Sequence 1011 BP; 305 A; 204 C; 213 G; 289 T; 0 U; 0 Other;
Query Match 52.1%; Score 702.2; DB 6; Length 1011;
Best Local Similarity 89.1%; Pred. NO. 7e-177;
Matches 758; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
497 TCGATGGTACTGTTGAAGTTAAAGAGGTGGATTGGAAGTTAACGGTCAATTTGTTAAAG 556
161 TCGACGGTACTGTTGAAGTTAAAGAGGTGGATTGGAAGTTAACGGTCAATTTGTTAAAG 220
557 TTTCTGCTGAACGGGAACAGCAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC 616
221 TTTCTGCTGAACGGGAACAGCAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC 280
617 TTGAAGCAACTAGTTCTTTGTTGTTAAAGAAAGCTGCTGAAACACATTACATGCTAACG 676
281 TTGAAGCAACTAGTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 340
677 GTGCTAAAGAAAGTTGTTATCACAGCTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 736
341 GTGCTAAAGAAAGTTGTTATCACAGCTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 400
737 ACACCTAACCCAGACATCTTGTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTA 796
401 ACACCTAACCCAGACATCTTGTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTA 460
797 CAACTGTTTGTAGCTCCTATGGCTAAAGCTCTTACAGATGATTTGGTATCCAAAAGGTC 856
461 CAACTGTTTGTAGCTCCTATGGCTAAAGCTCTTACAGATGATTTGGTATCCAAAAGGTC 520
857 TTATGACTACATCCACGCTTATATCTGTGACCAAAATGATCCTTGACGGACCAACCGTG 916
521 TGATGACTACTATCCACGCATACACTGTGACCAAAATGATCCTTGACGGACCAACCGTG 580
917 GTGGTGACCTTGTGCTGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
581 GTGGTGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640
977 CTGCTAAAGCTATCGGTTCTTGTATCCAGAAATGAAATGTTGTTGTTGTTGTTGTTGTTGTT 1036
641 CTGCTAAAGCTATCGGTTCTTGTATCCAGAAATGAAATGTTGTTGTTGTTGTTGTTGTTGTT 700
1037 AACGTGTTCTGTTCCAACTGGATCAGTAAGTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGG 1096
701 AACGTGTTCTGTTCCAACTGGATCAGTAAGTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGG 760
1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGGTTACA 1156

Db 761 TAACTGTGGAAGTAAGTAATGAGCTATGAAAGCAGCAGCTAACGATTTCATACGGTTATA 820
Qy 1157 CTGAAGATCCAATTTGTTCTTTCAGATATCGTAGGCGTGTATACGCGTTTCATTTGACG 1216
Db 821 CTGAAGATCCAATCGTATCATCTGATATCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 880
Qy 1217 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGTTAAAGTTGTTATCATGGT 1276
Db 881 CTACTCAAACTAAAGTTCAAACTGTTGACGGTAACCAATTTGTTAAAGTTGTTTCATGGT 940
Qy 1277 ATGACAATGAAATGTTTACACTGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1336
Db 941 ACGATAACGAAATGTTTACACTGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1000
Qy 1337 TCGCTAAATAA 1347
Db 1001 TCGCTAAATAA 1011
RESULT 10
ABA91249
ID ABA91249 standard; DNA; 1011 BP.
XX AC ABA91249;
XX DT 04-APR-2002 (first entry)
XX DE Streptococcus agalactiae gapC gene.
XX KW GapC; plasmin-binding protein; infection; mastitis; vaccine; diagnosis;
XX KM gene therapy; gene; ds.
XX OS Streptococcus agalactiae.
XX PN WO200196381-A2.
XX PD 20-DEC-2001.
XX PF 11-JUN-2001; 2001WO-CA000838.
XX PR 12-JUN-2000; 2000US-0211022P.
XX PA (UYSA-) UNIV SASKATCHEWAN.
XX PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
XX WPI; 2002-130725/17.
XX P-PSDB; AAM50640.
PT Novel isolated GapC protein of Streptococcus dysgalactiae, S. agalactiae,
PT S. uberis, S. parauberis, or S. iniae, useful as vaccine component for
PT treating streptococcal infection which causes mastitis in vertebrates.
XX Example 1; Fig 2A-B; 107pp; English.
XX
CC The present sequence is that of the coding region of the GapC plasmin
CC binding protein gene of Streptococcus agalactiae ATCC 27541. The gene was
CC amplified from chromosomal DNA by PCR using the primers given in ABA91253
CC -54. The PCR product was cloned into expression vector pE15b, creating
CC plasmid pMF521c. The gene encodes a GapC plasmin binding protein (see
CC AAM50640) that is capable of eliciting an immune response in a
CC vertebrate. The invention provides the GapC genes and proteins of 5
CC Streptococcus species, as well as recombinant vectors, host cells and
CC vaccine compositions comprising GapC polynucleotides or proteins. The
CC vaccines are used to treat or prevent a bacterial infection, especially a
CC streptococcal infection, and mastitis in particular (claimed). A
CC polynucleotide encoding a GapC protein is used in a claimed method of
CC treating or preventing a bacterial infection, such as a streptococcal
CC infection, especially mastitis. S. agalactiae is a common pathogen
CC associated with mastitis in cattle, horse, sheep and goat. It also causes
CC septicemia, meningitis, bacteraemia, impetigo, arthritis, urinary tract
CC infections, abscesses, spontaneous abortion, etc

Db 29942 CTGAAGATCCAAATCGTATCATCTGATATATCGTTGGTATTTCATACGGTTTCATGTTTGATG 29883
QY 1217 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGTATCATGTT 1276
Db 29882 CTACTCAAACTAAAGTTCAAACTGTTGACGGTAACCAATTTGGTTAAAGTTGTATCATGTT 29823
QY 1277 ATGACAAATGAAATGCTTACACTGCTCAACTTGTTCGTACACATTGAGTATTTTGCAGAAA 1336
Db 29822 ACGATAACGAAATGTCATACACTTTCACAACTTGTTCGTACACATTGAGTATTTTGCAGAAA 29763
QY 1337 TCGCTAAATAA 1347
Db 29762 TTGCAAAATAA 29752

RESULT 12
ABN71389
ID ABN71389 standard; DNA; 1008 BP.
XX
AC ABN71389;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 10691.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus agalactiae.

XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.

XX
PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
DR P-PSDB; ABP30758.
XX
PT New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.

XX
PS Claim 7; Page 4178-4179; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins

XX
SQ Sequence 1008 BP; 306 A; 203 C; 211 G; 288 T; 0 U; 0 Other;
Query Match 51.7%; Score 696; DB 6; Length 1008;
Best Local Similarity 88.8%; Pred. No. 3.2e-175;
Matches 753; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY 497 TCGATGGTACTGTTGAAGTTAAAGAGGTGGATTTCGAAGTTAAACGGTCAATTTGTTAAAG 556
Db 161 TCGACGGTACTGTTGAAGTTAAAGAGGTGGATTTCGAAGTTAAACGGTCAATTTGTTAAAG 220
QY 557 TTTCTGCTGAACGCGAACCCAGCAAAACATTGACTGGGTACTGATGGCGTAGAAATCGTTC 616
Db 221 TTTCTGCTGAACGCGAACCCAGCAAAACATTGACTGGGTACTGATGGCGTAGAAATCGTTC 280
QY 617 TTGAAGCAAACTAGTTTCTTTGCTAAAGAAAGCTGCTGAAAGAAACACTTACATGCTAACG 676
Db 281 TTGAAGCAAACTAGTTTCTTTGCTAAAGAAAGCTGCTGAAAGAAACACTTACATGCTAACG 340
QY 677 GTGCTAAAAAAGTTGTTTATCACAGCTCCTGGTGGAAACGACGTTTAAACAGTTGTTTCA 736
Db 341 GTGCTAAAAAAGTTGTTTATCACAGCTCCTGGTGGAAACGACGTTTAAACAGTTGTTTCA 400
QY 737 ACCTAACCCAGACATTTCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTAATA 796
Db 401 ACCTAACCCAGATATCTTTGATGGAACACTGAAACAGTTATCTCAGGTGCTTCATGTAATA 460
QY 797 CAAACTGTTTAGCTCCTTAAAGCTCTTTCACGATGCTATTTGGTATCCAAAAAGGTC 856
Db 461 CAAACTGTTTAGCTCCTTAAAGCTCTTTCACGATGCTATTTGGTATCCAAAAAGGTC 520
QY 857 TTATGACTACAAATCCACGCTTATCTGCTGACCAATGATCTTTCACGGACCCACCGGTG 916
Db 521 TGATGACTACTATCCACGCTATACCTGGTGACCAATGATCTTTCACGGACCCACCGGTG 580
QY 917 GTGCTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 581 GTGCTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
QY 977 CTGCTAAAGCTATCGGTCTTGTATCCAGAAATGGAATGGAATGGAATGGAATGGAATGGAATG 1036
Db 641 CTGCTAAAGCTATCGGTCTTGTATCCAGAAATGGAATGGAATGGAATGGAATGGAATGGAATG 700
QY 1037 AACGTGTTCTGTTCCAACTGGATCAGTAACAGTTGGTTGTAACCTCTTGTATGATAAAACG 1096
Db 701 AACGTGTTCTGTTCCAACTGGATCAGTAACAGTTGGTTGTAACCTCTTGTATGATAAAACG 760
QY 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAAACGACAGTTTCGGTTACA 1156
Db 761 TAAGTGTGGAAGAGTAATGAGCTATGAAAGCAGCAGCTAACGATTTCATACGGTTATA 820
QY 1157 CTGAAGATCCAATTTGTTTCTTTCAGATATCGTAGGCGGTGTCATACGGTTTCATTGTTGACG 1216
Db 821 CTGAAGATCCAATCGTATCATCTGATATCGTTGGTATTTTCATACGGTTTCATTGTTGATG 880
QY 1217 CAAGTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGTTAAAGTTGATCATGCT 1276
Db 881 CTAGTCAAACTAAAGTTCAAACTGTTGACGGTAACCAATTTGGTTAAAGTTGTTTCATGCT 940
QY 1277 ATGACAAATGAAATGCTTTCACACTGCTCAACTTGTTCGTACACTTGTAGTATTTTGCAGAAA 1336
Db 941 ACGATAACGAAATGTCATACACTTTCACAACTTGTTCGTACACTTGTAGTATTTTGCAGAAA 1000
QY 1337 TCGCTAAA 1344
Db 1001 TTGCAAAA 1008

RESULT 13
AAS55858
ID AAS55858 standard; DNA; 1080 BP.
XX
AC AAS55858;

PA (UUSA-) UNIV SASKATCHEWAN.
PI Potter AA, Perez-Casal J, Fontaine M;
XX WPI; 2002-098051/13.
DR P-PSDB; ABA50669.
XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
PT parauberis, or S. iniae GapC protein useful for treating mastitis in
PT vertebrates.
XX Example 3; Fig 5A-B; 116pp; English.
PS
XX The present sequence is that of the coding region of the gene encoding
CC the GapC plasmid binding protein, iniaeGapC (see AAM50669), of
CC Streptococcus iniae strain 9117. The gene was obtained by PCR
CC amplification of chromosomal DNA using the primers given in ABA91333-34.
CC The invention relates to novel GapC multiple epitope fusion proteins that
CC comprise epitopes from 1 or more of Streptococcus dysgalactiae,
CC Streptococcus agalactiae, Streptococcus parauberis, Streptococcus uberis
CC and Streptococcus iniae (see AAM50665-69). A claimed example is PolyGap4
CC (see AAM50664). The multiple epitope fusion proteins are produced using
CC claimed host-vector systems and used in claimed vaccines for treating or
CC preventing a bacterial infection in a vertebrate, especially a
CC streptococcal infection, and particularly mastitis. They are also used in
CC claimed methods of detecting Streptococcus antibodies. The multiple
CC epitope proteins are capable of eliciting broad immunity against a
CC variety of streptococcal infections while minimising the number of
CC antigens present in the final formulation and concomitantly reducing
CC production costs
XX
SQ Sequence 1011 BP; 305 A; 191 C; 213 G; 302 T; 0 U; 0 Other;
Query Match 50.3%; Score 678.2; DB 6; Length 1011;
Best Local Similarity 87.3%; Pred. No. 1.8e-170;
Matches 743; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 497 TCGATGGTACTGTTGAAGCTTAAGAGGTTGATTCGAAGTTAACGGTCAATTTGTTAAAG 556
DB 161 TTGACGGTACAGTTGAAGTTAAAGATGGTGGATTTCGAAGTTAACGGAGCTTTGTTAAAG 220
QY 557 TTTCTGCTGAACCGGACCCAGCAACATTGACTGGGCTACTGATGGCGTAGAATCGTTC 616
DB 221 TTTCTGACAGACCGGACCCAGCAACATTGACTGGGCTACTGATGGTGTAGACATCGTTC 280
QY 617 TTGAAGCAACTAGTTTCTTTGCTTAAAGAAAGAGCTGCTGAAACACATTACATGCTAACG 676
DB 281 TTGAAGCAACAGGTTTCTTCGCTTCTAAAGCAGCTGCTGAACACACATTACACGCTAACG 340
QY 677 GTGCTAAAGAAAGTTGTTATACAGCTCCTGTTGGAACAGCAGTTAAACACAGTTGTTTCA 736
DB 341 GTGCAAGAAAGTTGTTATACAGCTCCTGTTGGAATGACGTTAAACACAGTTGTTTACA 400
QY 737 ACATAACCCAGCAGCATTTCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA 796
DB 401 ACATAACCCAGCAGCATTTCTTGATGGAAGTGAACAGTTATCTCAGGTGCTTCATGTACTA 460
QY 797 CAAAGTGTCTAGCTCCTATGGTAAAGCTCTTTCACGATGCAATTTGGTATCCAAAAGGTC 856
DB 461 CAAAGTGTCTAGCTCCTATGGTAAAGCATTACAAGATAAATTTGGTGTAAACAAAGGTT 520
QY 857 TTATGACTACAATCCACGCTTATGCTGTTGACCAATGATCTTGAACGACACACCGGTG 916
DB 521 TAATGACTACTATCCATGTTTACACTGGTGGTACCAATGTTCTTGACGGACACACCGGTG 580
QY 917 GTGCTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 976
DB 581 GTGCTGATCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 640
QY 977 CTGCTAAAGCTATCGGCTCTGTTATCCAGAAATGATGTTGATGTTGCTGCTGCAC 1036
DB 641 CTGCTAAAGCAATCGGCTCTGTTATCCAGAAATGATGTTGATGTTGCTGCTGCAC 700

QY 1037 AACGTGTTCTCTCCAACTGGATCAGTAAGTGGTGGTGTAACTCTTGTATAAAACG 1096
DB 701 AACGTGTTCTCTCCAACTGGATCAGTAAGTGGTGGTGTAACTCTTGTATAAAACG 760
QY 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACACGACAGTTTCGGTTACA 1156
DB 761 CTTCAAGTAGAAGAAATCAATGCGAGCTATGAAAGCAGCAGTAACGATTCATACGGTTACA 820
QY 1157 CTGAAGATCCAAATGTTTCTTTCAGATATCGTAGGGCTGTACATCGGTTTCATTTGACG 1216
DB 821 CTGAAGATGCTATCGTATCATCATCATATCGTAGGTATTTCTTACGGTTTCATTTGATG 880
QY 1217 CAACTCAAACTAAAGTTATGGAAGTTGAGGATCAGGATCACAATTTGGTTAAAGTTGATCATGGT 1276
DB 881 CTACTCAAACTAAAGTACAAACTGTTGATGGAATCAATTTGGTTAAAGTTGTTTCATGGT 940
QY 1277 ATGACAATGAATGTCTTTACACCTGCTCAACTTTGTCGTACACTTGAGTATTTTGCAAAA 1336
DB 941 ATGACAATGAATGTCTTTACACCTGCTCAACTTTGTCGTACTCTTGTAGTACTTTTGCAAAA 1000
QY 1337 TCGCTAAATAA 1347
DB 1001 TCGCTAAATAA 1011
RESULT 15
ABA91252
ID ABA91252 standard; DNA; 1011 BP.
XX
AC ABA91252;
XX
DT 04-APR-2002 (first entry)
XX
DE Streptococcus iniae gapC gene.
XX
KW GapC; plasmin-binding protein; infection; mastitis; vaccine; diagnosis;
KW gene therapy; gene; ds.
XX
OS Streptococcus iniae.
XX
PN WO200196381-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000838.
XX
PR 12-JUN-2000; 2000US-0211022P.
XX
PA (UUSA-) UNIV SASKATCHEWAN.
XX
PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
XX
DR WPI; 2002-130725/17.
DR
P-PSDB; AAM50643.
XX
PT Novel isolated GapC protein of Streptococcus dysgalactiae, S. agalactiae,
PT S. uberis, S. parauberis, or S. iniae, useful as vaccine component for
PT treating streptococcal infection which causes mastitis in vertebrates.
XX
PS Example 1; Fig 5A-B; 107pp; English.
XX
CC The present sequence is that of the coding region of the GapC plasmid
CC binding protein gene of Streptococcus iniae strain 9117. The gene was
CC amplified from chromosomal DNA by PCR using the primers given in ABA91253
CC and ABA91254. The PCR product was cloned into expression vector pET15b,
CC creating plasmid pMF521e. The gene encodes a GapC plasmid binding protein
CC (see AAM50643) that is capable of eliciting an immune response in a
CC vertebrate. The invention provides the GapC genes and proteins of 5
CC Streptococcus species, as well as recombinant vectors, host cells and
CC vaccine compositions comprising GapC polynucleotides or proteins. The
CC vaccines are used to treat or prevent a bacterial infection, especially a
CC streptococcal infection, and mastitis in particular (claimed). A

CC polynucleotide encoding a GapC protein is used in a claimed method of
CC treating or preventing a bacterial infection, such as a streptococcal
CC infection, especially mastitis
XX

SQ Sequence 1011 BP; 305 A; 191 C; 213 G; 302 T; 0 U; 0 Other;
Query Match 50.3%; Score 678.2; DB 6; Length 1011;
Best Local Similarity 87.3%; Pred. No. 1.8e-170;
Matches 743; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY	497	TCGATGGTACTGTTGAAGTTAAAGAAGGTGGATTTCGAAGTTAACGGTCAATTTGTTAAAG	556
Db	161	TTGACGTACAGTTGAAGTTAAAGATGGTGGATTTCGAAGTTAACGGAAGCTTTGTTAAAG	220
QY	557	TTTCTGCTGAACGCGAACCGCAAAACATTGACTGGCTACTGATGGCGTAGAATCGTTTC	616
Db	221	TTTCTGCAGAACGCGAACCGCAAAACATTGACTGGCTACTGATGGGTAGACATCGTTTC	280
QY	617	TTGAAGCAACTAGTTTCTTTGCTTAAAGAAAGAGCTGCTGAAAAACACATTACATGCTAACG	676
Db	281	TTGAAGCAACAGGTTTCTTCGCTTCTAAAGCAGCTGCTGAAACACACATTACGCTAACG	340
QY	677	GTGCTAAAAAAGTTGTTATACACAGCTCCTGGTGGAAACGACGTTAAAAACAGTTGTTTCA	736
Db	341	GTGCGAAAAAAGTTGTTATACACAGCTCCTGGTGGAAATGACGTTAAAAACAGTTGTTTACA	400
QY	737	ACACTAACCCACGACATTCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA	796
Db	401	ACACTAACCATGATATCTTTGATGGAACCTGAAACAGTTATCTCAGGTGCTTCATGTACTA	460
QY	797	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTCACGATGCATTTGGTATCCAAAAAGGTC	856
Db	461	CAAACTGTTTAGCTCCTAATGGCTAAAGCATTTACAGATAAATCTTTGGTGTAAACAAGGTT	520
QY	857	TTATGACTACAAATCCACGCTTATCTGTTGAGCCAAATGATCCTTGACGGACACACCCGTG	916
Db	521	TAATGACTACTATCCATGGTTTACACTGGTGACCAAAATGGTTCTTTGACGGACACACCCGTG	580
QY	917	GTGGTGACCTTCGTGCTGCTGCTGGTGCTGCAAAACATTTGTTCTTAACCTCAACTGGTG	976
Db	581	GTGGTGATCTTCGTGCTGCTGCTGCTGCAAGCAAAACATCGTTCTTAACCTCAACTGGTG	640
QY	977	CTGCTAAAGCTATCGGTCTTGTATCCAGAAATGAATGGTAAACCTTGATGGTGTGTCAC	1036
Db	641	CTGCTAAAGCAATCGGTCTTGTATCCAGAAATGAATGGTAAACCTTGACGGTGTGTCAC	700
QY	1037	AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACTCTTGATAAAAAACG	1096
Db	701	AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGTAACTCTTGATAAAAAAGATA	760
QY	1097	TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAAACGACAGTTTCGGTTACA	1156
Db	761	CTTCAGTAGAAGAAATCAATGCGCTATGAAAGCAGCAGCTAACGATTTCATACGGTTACA	820
QY	1157	CTGAAGATCCAAATGTTTCTTCAGATATCGTAGGGGTGTACATACGGTTTCAATGTTGACG	1216
Db	821	CTGAAGATGCTATCGTATCATCAGATATCGTAGGTATTTCTTACGGTTCAATTTTGTATG	880
QY	1217	CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGTATCATGGT	1276
Db	881	CTACTCAAACTAAAGTACAAAACCTGTTGATGGAATCAATTTGGTTAAAGTTGTTTCATGGT	940
QY	1277	ATGACAATGAAATGCTTACACTGCTCAACTGTTGTCGACACTTGAGTATTTTGCAAAA	1336
Db	941	ATGACAATGAAATGCTTACACTGCTCAACTGTTGTCGACTCTTTGAGTACTTTTGCAAAA	1000
QY	1337	TCGCTAAATAA 1347	
Db	1001	TCGCTAAATAA 1011	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 23:30:10 ; Search time 128 Seconds
(without alignments)
5839.991 Million cell updates/sec

Title: US-10-650-369-21

Perfect score: 1347

Sequence: 1 atgaaaaaaataacaggat.....ttgcaaaaatcgctaaataa 1347

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A-COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B-COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A-COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B-COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS-COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1347	100.0	1347	4	US-09-878-766A-21	Sequence 21, Appl
2	815.8	60.6	1011	4	US-09-878-766A-11	Sequence 11, Appl
3	780.6	58.0	1125	1	US-07-928-462-1	Sequence 1, Appl
4	780.6	58.0	1125	3	US-08-273-247-1	Sequence 1, Appl
5	702.2	52.1	1011	4	US-09-878-766A-13	Sequence 13, Appl
6	678.2	50.3	1011	4	US-09-878-766A-19	Sequence 19, Appl
7	678.2	50.3	7571	4	US-08-961-527-2	Sequence 2, Appl
8	676.6	50.2	1011	4	US-09-878-766A-15	Sequence 15, Appl
9	660.2	49.0	1000	3	US-08-961-083-53	Sequence 53, Appl
10	660.2	49.0	1000	4	US-09-536-784-53	Sequence 53, Appl
11	639.8	47.5	1011	4	US-09-878-766A-17	Sequence 17, Appl
12	495.6	36.8	1002	4	US-09-107-532A-1115	Sequence 1115, Ap
13	478	35.5	1041	4	US-09-134-000C-995	Sequence 995, App
14	426.8	31.7	1047	4	US-09-134-001C-2676	Sequence 2676, Ap
15	409.2	30.4	6048	4	US-08-956-171E-128	Sequence 128, App
16	277	20.6	1074	4	US-09-134-000C-824	Sequence 824, App
17	269	20.0	849	4	US-09-107-532A-403	Sequence 403, App
18	268.2	19.9	580073	4	US-08-545-528D-1	Sequence 1, Appl
19	239.4	17.8	1362	6	5290690-8	Patent No. 5290690
20	235.6	17.5	1717	4	US-09-634-238-18	Sequence 18, Appl
21	221.8	16.5	1928	4	US-09-674-826B-5	Sequence 5, Appl
22	210	15.6	1023	4	US-09-543-681A-3911	Sequence 3911, Ap
23	205	15.2	1050	4	US-09-489-039A-6768	Sequence 6768, Ap
24	198.6	14.7	1008	4	US-09-489-039A-508	Sequence 508, App
25	177.2	13.2	1230025	4	US-09-198-452A-1	Sequence 1, Appl
26	169.8	12.6	35881	4	US-08-311-731A-127	Sequence 127, App
27	169.2	12.6	1026	4	US-09-134-001C-946	Sequence 946, App

28	160.4	11.9	4409	4	US-09-653-403-2	Sequence 2, Appli
29	160.4	11.9	4409	4	US-10-013-784-2	Sequence 2, Appli
30	160.4	11.9	4409	4	US-09-786-606-1	Sequence 1, Appli
31	160.4	11.9	4431	3	US-09-532-803-8	Sequence 8, Appli
32	159.2	11.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
33	159.2	11.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
34	154	11.4	3333	3	US-09-532-803-10	Sequence 10, Appli
35	154	11.4	3333	4	US-09-653-403-1	Sequence 1, Appli
36	154	11.4	3333	4	US-10-013-784-1	Sequence 1, Appli
37	151	11.2	1364	3	US-09-095-855-204	Sequence 204, App
38	151	11.2	1364	4	US-09-205-426-204	Sequence 204, App
39	145	10.8	2375	2	US-08-903-800A-5	Sequence 5, Appli
40	144.2	10.7	3152	4	US-09-221-017B-931	Sequence 931, App
41	144	10.7	1830121	4	US-09-557-884-1	Sequence 1, Appli
42	144	10.7	1830121	4	US-09-643-990A-1	Sequence 1, Appli
43	142	10.5	2574	4	US-08-956-171E-441	Sequence 441, App
44	140.6	10.4	898	2	US-08-997-080-185	Sequence 185, App
45	140.6	10.4	898	2	US-08-997-362-185	Sequence 185, App

ALIGNMENTS

RESULT 1
US-09-878-766A-21
; Sequence 21, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer Gap4
; OTHER INFORMATION: chimeric GapC protein
; NAME/KEY: CDS
; LOCATION: (1)..(1347)
US-09-878-766A-21

Query Match	100.0%;	Score 1347;	DB 4;	Length 1347;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1347;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATGAAAAAAATAACAGGGATTATTTTATTGCTTCTTTCAGTCATTATTCTGTCATGC	60	
Db	1	ATGAAAAAAATAACAGGGATTATTTTATTGCTTCTTTCAGTCATTATTCTGTCATGC	60	
Qy	61	CAGGCAAACTACGGATCCGGTATGGTAGTAAAGTTGGTATTAAACGGTTTCGGTCGTATC	120	
Db	61	CAGGCAAACTACGGATCCGGTATGGTAGTAAAGTTGGTATTAAACGGTTTCGGTCGTATC	120	
Qy	121	GGACGCTTTCGATCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC	180	
Db	121	GGACGCTTTCGATCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC	180	
Qy	181	GACCTTACAGATCCAAACATGCTTGCACACTTGTGAAATACGATACAACTCAAGGACGT	240	
Db	181	GACCTTACAGATCCAAACATGCTTGCACACTTGTGAAATACGATACAACTCAAGGACGT	240	
Qy	241	TTTGACGGAACTGTTGAAGTTAAAGAGGTGGATTGTAAGTAACGAACTTCATCAAA	300	
Db	241	TTTGACGGAACTGTTGAAGTTAAAGAGGTGGATTGTAAGTAACGAACTTCATCAAA	300	

QY	301	GT	TTCTGCTGAACG	TGATCCAG	AAAAACATCG	ACTG	GGCAACTG	ACGGTGT	TGAAATCG	TT	360
DB	301	GT	TTCTGCTGAACG	TGATCCAG	AAAAACATCG	ACTG	GGCAACTG	ACGGTGT	TGAAATCG	TT	360
QY	361	CT	GGAAGCACTCG	AGGGTACT	GTAGAAG	TAAAGAT	GGTGGAT	TTTGAC	TTTGAAC	GGAAAA	420
DB	361	CT	GGAAGCACTCG	AGGGTACT	GTAGAAG	TAAAGAT	GGTGGAT	TTTGAC	TTTGAAC	GGAAAA	420
QY	421	TT	CATTAAGTTTCT	GTGTA	AAAAAGATCC	AGAACAA	ATTGACT	GGCAACTG	ACGGTGT	TT	480
DB	421	TT	CATTAAGTTTCT	GTGTA	AAAAAGATCC	AGAACAA	ATTGACT	GGCAACTG	ACGGTGT	TT	480
QY	481	GA	AATCGTTCTT	GAAATCG	ATGGTACT	GTGTAAG	TAAAGAG	STGATTC	GAAATTAAC		540
DB	481	GA	AATCGTTCTT	GAAATCG	ATGGTACT	GTGTAAG	TAAAGAG	STGATTC	GAAATTAAC		540
QY	541	GG	TCAATTTGTT	ATAAGTTTCT	GCTGAACG	GGAACCA	CCAGCAAC	ATTGACT	GGCTACT	GTAT	600
DB	541	GG	TCAATTTGTT	ATAAGTTTCT	GCTGAACG	GGAACCA	CCAGCAAC	ATTGACT	GGCTACT	GTAT	600
QY	601	GG	CGTAGAAATCG	TTTCTTGA	AGCAACTAG	TTTCTTTG	CTATA	AAAAAGAA	AGCTGCT	GAAAAA	660
DB	601	GG	CGTAGAAATCG	TTTCTTGA	AGCAACTAG	TTTCTTTG	CTATA	AAAAAGAA	AGCTGCT	GAAAAA	660
QY	661	CA	CTTACATGCT	TAACG	STGCTAAAA	AAGTTGTT	TATCAC	AGCTCCT	GGTGGAA	ACGACG	720
DB	661	CA	CTTACATGCT	TAACG	STGCTAAAA	AAGTTGTT	TATCAC	AGCTCCT	GGTGGAA	ACGACG	720
QY	721	AA	ACAGTTGTTT	TTCAAC	ACTAACC	ACGACAT	TTCTTG	ACGGTACT	GAAACAG	TTATCTCA	780
DB	721	AA	ACAGTTGTTT	TTCAAC	ACTAACC	ACGACAT	TTCTTG	ACGGTACT	GAAACAG	TTATCTCA	780
QY	781	GG	TGCTTCATG	TACTAC	AAACTGTT	TAGCTCCT	TATGG	CTAAAGCT	CTTCA	CGATGCAT	840
DB	781	GG	TGCTTCATG	TACTAC	AAACTGTT	TAGCTCCT	TATGG	CTAAAGCT	CTTCA	CGATGCAT	840
QY	841	GG	TATCCAAAAAG	STCTTAT	GACTACA	ATCCAC	GCCTTAT	ACTGGTG	ACCAAT	GATCCT	900
DB	841	GG	TATCCAAAAAG	STCTTAT	GACTACA	ATCCAC	GCCTTAT	ACTGGTG	ACCAAT	GATCCT	900
QY	901	GA	CGGACCA	CACCGTGG	TGGTGAC	CTTCG	TGCTGCT	CGTGGT	GCTGCA	AAACAT	960
DB	901	GA	CGGACCA	CACCGTGG	TGGTGAC	CTTCG	TGCTGCT	CGTGGT	GCTGCA	AAACAT	960
QY	961	CC	TAACTCAACT	GGTGTCT	GCTAAAG	CTATCG	GTGCT	CGTGGT	GCTGCA	AAACAT	1020
DB	961	CC	TAACTCAACT	GGTGTCT	GCTAAAG	CTATCG	GTGCT	CGTGGT	GCTGCA	AAACAT	1020
QY	1021	CT	TGATGGT	GTCACAA	CGTGTTC	CTGTTCC	AACTGG	ATCCAG	AATTCG	TGTA	1080
DB	1021	CT	TGATGGT	GTCACAA	CGTGTTC	CTGTTCC	AACTGG	ATCCAG	AATTCG	TGTA	1080
QY	1081	AC	TCTTGATA	AAAAACG	TTTCTG	TTGAC	GAATCA	ACGCTG	CTATG	AAAGCTG	1140
DB	1081	AC	TCTTGATA	AAAAACG	TTTCTG	TTGAC	GAATCA	ACGCTG	CTATG	AAAGCTG	1140
QY	1141	GA	CAGTTTCG	GTACACT	GGAAGATCC	AAATTG	TTTCTTC	CAGATAT	CGTAG	GGGTGCAT	1200
DB	1141	GA	CAGTTTCG	GTACACT	GGAAGATCC	AAATTG	TTTCTTC	CAGATAT	CGTAG	GGGTGCAT	1200
QY	1201	GG	TTCA	TTGTTTG	ACGCAACT	CAAACTAA	AGTTAT	GGAAGTT	GAC	CGGATC	1260
DB	1201	GG	TTCA	TTGTTTG	ACGCAACT	CAAACTAA	AGTTAT	GGAAGTT	GAC	CGGATC	1260
QY	1261	AA	AGTTG	TATCAT	GGTATG	ACAATG	AAATG	TCTTAC	ACTGCT	CAACTT	1320
DB	1261	AA	AGTTG	TATCAT	GGTATG	ACAATG	AAATG	TCTTAC	ACTGCT	CAACTT	1320
QY	1321	GAG	TATTTTG	CAAAAAAT	CGCTAA	ATAA					1347
DB	1321	GAG	TATTTTG	CAAAAAAT	CGCTAA	ATAA					1347

```

RESULT 2
US-09-878-766A-11
; Sequence 11, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE AGAINST STREPTOCOCCUS
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus dysgalactiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1011)
US-09-878-766A-11

```

Query Match 60.6%; Score 815.8; DB 4; Length 1011;
Best Local Similarity 97.4%; Pred. No. 8.1e-215;
Matches 829; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

iv


```
Db 821 CTGAAGATCCAAATGTTCTTCAGATATCGTAGGCGTGTCATACGGTTTCATTGTTGACG 880
QY 1217 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCAACAATGTTAAAGTTGTATCATGGT 1276
Db 881 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCAACAATGTTAAAGTTGTATCATGGT 940
QY 1277 ATGACAAATGAAATGTCTTACACTGCTCAACTTGTTCGTACACTTGAGTATTTTGCAAAA 1336
Db 941 ATGACAAATGAAATGTCTTACACTGCTCAACTTGTTCGTACACTTGAGTATTTTGCAAAA 1000
QY 1337 TCGCTAAATAA 1347
Db 1001 TCGCTAAATAA 1011
```

```
RESULT 3
US-07-928-462-1
; Sequence 1, Application US/07928462
; Patent No. 5328996
; GENERAL INFORMATION:
; APPLICANT: Boyle, Michael D.P.
; APPLICANT: Lottenberg, Richard
; APPLICANT: Broder, Christopher C.
; APPLICANT: von Mering, Gregory O.
; TITLE OF INVENTION: Bacterial Plasmin Receptors as
; TITLE OF INVENTION: Fibrinolytic Agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,462
; FILING DATE: 19920810
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/524,411
; FILING DATE: 16-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/330,849
; FILING DATE: 29-MAR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1125 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; STRAIN: M untypable
; INDIVIDUAL ISOLATE: 64/14
; IMMEDIATE SOURCE:
; CLONE: pRL015
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: 115..1122
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 115..1122
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /codon_start= 115
; OTHER INFORMATION: /function= "High-affinity binding of
; OTHER INFORMATION: plasmin(ogen)"
; OTHER INFORMATION: /product= "Streptococcal plasmin receptor"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "plr"
; OTHER INFORMATION: /number= 1
; OTHER INFORMATION: /label= PLR
US-07-928-462-1
```

```
Query Match 58.0%; Score 780.6; DB 1; Length 1125;
Best Local Similarity 94.8%; Pred. No. 4.2e-205;
Matches 807; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 497 TCGATGGTACTGTTGAAGTTAAAGAGGTGGATTGGAAGTTAAACGGTCAATTTGTTAAAG 556
Db 275 TTGATGGAACAGTTGAAGTTAAAGAGGTGGATTGGAAGTTAAACGGTCAATTTGTTAAAG 334
QY 557 TTTCTGCTGAACGGAACAGCAAACTATGACTGGGCTACTGATGGCTAGAAATCGTTC 616
Db 335 TTTCTGCTGAACGATCCAGAAACATCGACTGGGCAACTGATGGGTTGAAATCGTTC 394
QY 617 TTGAAGCAACTAGTTTCTTTGCTAAAGAAAGCTGCTGAAACACATTACATGCTAACG 676
Db 395 TTGAAGCAACTGGTTTCTTTGCTAAAGAAAGCTGCTGAAACACATTACATGCTAACG 454
QY 677 GTGCTAAAAAGTTGTTTATCACAGCTCCTGGTGAAACGACGTTAAACAGTTGTTTCA 736
Db 455 GTGCTAAAAAGTTGTTTATCACAGCTCCTGGTGAAACGATGTTAAACAGTTGTTTCA 514
QY 737 AACTAACCCAGACATTCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCACTACTA 796
Db 515 AACTAACCCAGACATTCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCACTACTA 574
QY 797 CAAACTGTTTAGTCTCTATGGCTAAAGCTCTTACGATGCAATTCGATGTTCAAAAGGTC 856
Db 575 CAAACTGTTTAGTCTCTATGGCTAAAGCTCTTACGATGCAATTCGATGTTCAAAAGGTC 634
QY 857 TTATGACTACAAATCCACGCTTATCTGCTGACCAATGATCCTTGACGGACCAACCGTG 916
Db 635 TTATGACTACAAATCCACGCTTATCTGCTGACCAATGATCCTTGACGGACCAACCGTG 694
QY 917 GTGGTGACCTTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 695 GTGGTGACCTTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 754
QY 977 CTGCTAAAGCTATCGGTCTTCTGTTATCCAGAAATGGAATGGTAAACTTGTGCTGCTGCT 1036
Db 755 CTGCTAAAGCTATCGGTCTTCTGTTATCCAGAAATGGAATGGTAAACTTGTGCTGCTGCT 814
QY 1037 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACCTCTTGATAAAACG 1096
Db 815 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACCTCTTGATAAAACG 874
QY 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGGTTACA 1156
Db 875 TTTCTGTTGACGAAATCAACCTGCTATGAAAGCTGCTTCAACGACAGTTTCGGTTACA 934
QY 1157 CTGAAGATCCAAATGTTTCTTTCAGATATCGTAGGCGTGTATACGTTTCATTGTTGACG 1216
Db 935 CTGAAGATCCAAATGTTTCTTTCAGATATCGTAGGCGTGTATACGTTTCATTGTTGACG 994
QY 1217 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCAACAATGTTAAAGTTGTATCATGGT 1276
Db 995 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCAACAATGTTAAAGTTGTATCATGGT 1054
QY 1277 ATGACAAATGAAATGTCTTACACTGCTCAACTGTTTCTGCTGCTGCTGCTGCTGCTGCTG 1336
```

Db 1055 ATGACAAACGAATGTCTTACACTGCTCAACTGTACGTAAGTCTGAGTACTTCGCAAAAA 1114
QY 1337 TCGCTAAATAA 1347
Db 1115 TTGCTAAATAA 1125

RESULT 4
US-08-273-247-1
; Sequence 1, Application US/08273247
; Patent No. 6136323
; GENERAL INFORMATION:
; APPLICANT: Boyle, Michael D.P.
; APPLICANT: Lottenberg, Richard
; APPLICANT: Broder, Christopher C.
; APPLICANT: von Mering, Gregory O.
; TITLE OF INVENTION: Bacterial Plasmin Receptors as
; TITLE OF INVENTION: Fibrinolytic Agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,247
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/524,411
; FILING DATE: 16-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/330,849
; FILING DATE: 29-MAR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1125 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; STRAIN: M untypable
; INDIVIDUAL ISOLATE: 64/14
; IMMEDIATE SOURCE:
; CLONE: PRL015
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 115..1122
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 115..1122
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /codon start= 115
; OTHER INFORMATION: /function= "High-affinity binding of

; OTHER INFORMATION: plasmin(ogen)"
; OTHER INFORMATION: /product= "Streptococcal plasmin receptor"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "plr"
; OTHER INFORMATION: /number= 1
; OTHER INFORMATION: /label= PLR
US-08-273-247-1

Query Match 58.0%; Score 780.6; DB 3; Length 1125;
Best Local Similarity 94.8%; Pred. No. 4.2e-205;
Matches 807; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 497 TCGATGGTACTGTGAAGTTAAAGAGGTGATTCGAAGTTAAACGGTCAATTGTTAAAG 556
Db 275 TTGATGGAACAGTTGAAGTTAAAGAGGTGATTTGAAGTAAACGGAAACTTCATCAAAG 334
QY 557 TTTCTGCTGAACGGCAACAGCAACATTCAGTGGGCTACTGATGGCGTAGAAATCGTTC 616
Db 335 TTTCTGCTGAACGGTATCCAGAAACATCGACTGGGCAACTGATGGGTTGAAATCGTTC 394
QY 617 TTGAAGCAACTAGTTTCTTTGCTAAAAAAGAGCTGCTGAAAAACACTTACATGCTAACG 676
Db 395 TTGAAGCAACTGGTTTCTTTGCTAAAAAAGAGAGCTGAAAAACACTTACATGCTAACG 454
QY 677 GTGCTAAAAAGTTGTTATCACAGCTCCTGGTGGAAACAGCACTTTGTTTCA 736
Db 455 GTGCTAAAAAGTTGTTATCACAGCTCCTGGTGGAAACAGATGTTAAACAGTTGTTTCA 514
QY 737 ACATAACCAACAGACATCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCACTACTA 796
Db 515 ACATAACCAACAGACATCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCACTACTA 574
QY 797 CAAACTGTTTAGTCTCCTATGGCTAAAGCTCTTACGATGATGATTTGGTATCCAAAAAGTC 856
Db 575 CAAACTGTTTAGTCTCCTATGGCTAAAGCTCTTACGATGATTTGGTATCCAAAAAGTC 634
QY 857 TTATGACTACATCCACGCTTATCTGAGTACCAAAATGATCCTTGACGGACCAACCGTG 916
Db 635 TTATGACTACATCCACGCTTATCTGAGTACCAAAATGATCCTTGACGGACCAACCGTG 694
QY 917 GTGGTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 976
Db 695 GTGGTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 754
QY 977 CTGCTAAAGCTATCGTCTTGTATCCAGAAATGAATGGTAAACCTTCACTGCTGCAC 1036
Db 755 CTGCTAAAGCTATCGTCTTGTATCCAGAAATGAATGGTAAACCTTCACTGCTGCAC 814
QY 1037 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACCTTTGATAAAAAACG 1096
Db 815 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACCTTTGATAAAAAACG 874
QY 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAAACGACAGTTTCGGTTACA 1156
Db 875 TTTCTGTTGACGAAATCAACCTGCTATGAAAGCTGCTTCAAAACGATAGCTTCGGTTACA 934
QY 1157 CTGAAGATCCAATTTCTTTCAGATATCGTAGGCGTGTATACAGGTTCACTGTTTACG 1216
Db 935 CTGAAGATCCAATCGTTTCTTTCAGATATCGTAGGCGTGTATACAGGTTCACTGTTTACG 994
QY 1217 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGATCATGGT 1276
Db 995 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGATCATGGT 1054
QY 1277 ATGACAATGAAATGTTTACACTGCTCAACTTGTTCGTACACTTTGAGTATTTGCAAAAA 1336
Db 1055 ATGACAACGAAATGTTTACACTGCTCAACTTGTTCGTACACTTTGAGTACTTTCGCAAAAA 1114
QY 1337 TCGCTAAATAA 1347
Db 1115 TTGCTAAATAA 1125

RESULT 5
US-09-878-766A-13
; Sequence 13, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1011)
US-09-878-766A-13

Query Match 52.1%; Score 702.2; DB 4; Length 1011;
Best Local Similarity 89.1%; Pred. No. 1.5e-183;
Matches 758; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY	497	TCGATGGTACTGTTGAAGTTAAAGAGGTGGATTTCGAAGTTAACGGTCAATTGTTAAAG	556
Db	161	TCGACGGTACTGTTGAAGTTAAAGAGGTGGATTTCGAAGTTAACGGTCAATTGTTAAAG	220
QY	557	TTTCTGCTGAACGCGAACCCAGCAAAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC	616
Db	221	TTTCTGCTGAACGCGAACCCAGCAAAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC	280
QY	617	TTGAAGCAACTAGTTTCTTTGCTTAAAGAAAGAGCTGCTGAAAAACACTTACATGCTAACG	676
Db	281	TTGAAGCAACTAGTTTCTTTGCTTAAAGAAAGAGCTGCTGAAAAACACTTACATGCTAACG	340
QY	677	GTGCTAAAAAGTTGTTATCACAGCTCCTGTTGGAACCGAGCTTAAACAGTTGTTTCA	736
Db	341	GTGCTAAAAAGTTGTTATCACAGCTCCTGTTGGAACCGAGCTTAAACAGTTGTTTCA	400
QY	737	ACACTAACCCAGCATATCCTTGATGGAACAGTTCATCAGGTGCTTCATGCTACTA	796
Db	401	ACACTAACCCAGCATATCCTTGATGGAACAGTTCATCAGGTGCTTCATGCTACTA	460
QY	797	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTCACGATGCATTTGGTATCCAAAAAGGTC	856
Db	461	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTCACGATGCATTTGGTATCCAAAAAGGTC	520
QY	857	TTATGACTACAAATCCAGCTTATACAGCTCCTGTTGGAACCGAGCTTAAACAGTTGTTTCA	916
Db	521	TGATGACTACAAATCCAGCTTATACAGCTCCTGTTGGAACCGAGCTTAAACAGTTGTTTCA	580
QY	917	GTGCTAACCCAGCATATCCTTGATGGAACAGTTCATCAGGTGCTTCATGCTACTA	976
Db	581	GTGCTAACCCAGCATATCCTTGATGGAACAGTTCATCAGGTGCTTCATGCTACTA	640
QY	977	CTGCTAAAGCTATCGGTCTGTTATCCAGAAATGAAATGGTAAAGTTCATGCTGCTGAC	1036
Db	641	CTGCTAAAGCTATCGGTCTGTTATCCAGAAATGAAATGGTAAAGTTCATGCTGCTGAC	700
QY	1037	AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTGTAACTCTTGATAAAACG	1096
Db	701	AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTGTAACTCTTGATAAAACG	760
QY	1097	TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGGTTACA	1156
Db	761	TAACTGTCGAAGAAGTAAATGCAGCTATGAAAGCAGCAGCTAACGATTTCATACGGTTATA	820
QY	1157	CTGAAGATCCAATGTTTCTTCAGATATCGTAGGCGTGTCAACGTTTCATGTTTGACG	1216

Db	821	CTGAAGATCCAATCGTATCATCTGATATCGTTGGTATTTTCATACGGTTCATTTGTTGATG	880
QY	1217	CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTCGTTAAAGTTGTATCATGGT	1276
Db	881	CTACTCAAACTAAAGTTCAAACTGTTGACGGTAACCAATTCGTTAAAGTTGTTTCATGGT	940
QY	1277	ATGACAAATGAAATGTCTTACACTGCTCAACTTGTTCGTACACTTGAGTATTTTGCAAAAA	1336
Db	941	ACGATAACGAAATGTCACTACACTTCACAACTTGTTCGTACACTTGAGTACTTTTGCAAAAA	1000
QY	1337	TCGCTAAATAA 1347	
Db	1001	TCGCTAAATAA 1011	

RESULT 6

US-09-878-766A-19
; Sequence 19, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus iniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1011)
US-09-878-766A-19

Query Match 50.3%; Score 678.2; DB 4; Length 1011;
Best Local Similarity 87.3%; Pred. No. 6e-177;
Matches 743; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY	497	TCGATGGTACTGTTGAAGTTAAAGAGGTGGATTTCGAAGTTAACGGTCAATTGTTAAAG	556
Db	161	TTGACGGTACAGTTGAAGTTAAAGAGGTGGATTTCGAAGTTAACGGTCAATTGTTAAAG	220
QY	557	TTTCTGCTGAACGCGAACCCAGCAAAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC	616
Db	221	TTTCTGCTGAACGCGAACCCAGCAAAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC	280
QY	617	TTGAAGCAACTAGTTTCTTTGCTTAAAGAAAGAGCTGCTGAAAAACACTTACATGCTAACG	676
Db	281	TTGAAGCAACTAGTTTCTTTGCTTAAAGAAAGAGCTGCTGAAAAACACTTACATGCTAACG	340
QY	677	GTGCTAAAAAGTTGTTATCACAGCTCCTGTTGGAACCGAGCTTAAACAGTTGTTTCA	736
Db	341	GTGCTAAAAAGTTGTTATCACAGCTCCTGTTGGAACCGAGCTTAAACAGTTGTTTCA	400
QY	737	ACACTAACCCAGCATATCCTTGATGGAACAGTTCATCAGGTGCTTCATGCTACTA	796
Db	401	ACACTAACCCAGCATATCCTTGATGGAACAGTTCATCAGGTGCTTCATGCTACTA	460
QY	797	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTCACGATGCATTTGGTATCCAAAAAGGTC	856
Db	461	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTCACGATGCATTTGGTATCCAAAAAGGTC	520
QY	857	TTATGACTACAAATCCAGCTTATACAGCTCCTGTTGGAACCGAGCTTAAACAGTTGTTTCA	916
Db	521	TAACTGACTACTATCCATGGTTACACTGGTGACCAATGGTTCTTGACGGACCAACCGTG	580
QY	917	GTGCTAACCCAGCTTCGTCGTCTGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	976

396	ACACTAACCCACGACGTTCTTTGACGGTACTGAAACAGTTTATCTCAGGTGCTTTCATGTACTA	455
797	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTTCACGATGCAATTTGGTATCCAAAAAGGTC	856
456	CAAACTGCTTGGCTCCAATGGCTAAAGCTCTTCAAGACAACTTTGGTGTGTTGAAGGAT	515
857	TTATGACTACAATCCACGCTTATACTGGTGACCAAAATGATCCTTTGACGGACCACACCGTG	916
516	TGATGACTACTATCCACGCTTACACTGGTGACCAAAATGATCCTTTGACGGACCACACCGTG	575
917	GTGGTGACCTTCGTCGTGCTCGTCTGGTGTGCTGCAAAATGTTCCCTAACTCAACTGGTG	976
576	GTGGTGACCTTCGCCGCTGCTGGCGCTGGTGTGCAAAATGTTCCCTAACTCAACTGGTG	635
977	CTGCTAAAGCTATCGGTCTTGTATTATCCAGAAATTGAATGGTAAACTTGTATGGTGTGTCAC	1036
636	CTGCAAAAGCTATCGGTCTTGTAAATCCAGAAATTGAATGGTAAACTTGTACGGATCTGCAC	695
1037	AACGTGTTCTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACTCTTGATAAAAAACG	1096
696	AACGCGTTCCAACTCCAACTGGATCAGTTACTGAATGGTAGCAGTTCTTGAAGAAACG	755
1097	TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAAACGACAGTTTCGGTTACA	1156
756	TTACTGTTGATGAAGTGAACGCAGCTATGAAAGCAGCTTCAAAACGAATCATACGGTTACA	815
1157	CTGAAGATCCAATTGTTTCTTCAGATATCGTAGGCGGTGTATACGGTTCAATTGTTGACG	1216
816	CAGAAGATCCAATCGTATCTTCAGATATCGTAGGTATGTCTTACGGTTCAATTGTTGACG	875
1217	CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTGGTTAAAGTTGTATCATGGT	1276
876	CAACTCAAACTAAAGTTCTTGACGTTGACGGTAAACAAATTGGTTAAAGTTGTATCATGGT	935
1277	ATGACAATGAAATGTCTTACACTGCTCAACTTGTGTCGTACACTT-GAGTATTTTGCAAAA	1335
936	ACGACAAACGAAATGTATACACTGCACTGCACTTGTTCGTACTCTTTGGAATACTTCGCAAAA	995
1336	ATCGC 1340	
996	ATTGC 1000	

RESULT 10

US-09-536-784-53
; Sequence 53, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971

	Query Match	36.8%;	Score 495.6;	DB 4;	Length 1002;
	Best Local Similarity	75.3%;	Pred. No. 1.1e-126;		
	Matches 631;	Conservative	0;	Mismatches 204;	Indels 3; Gaps 1;
QY	497	TCGATGGTACTGTTGAAGTTAAAGAAAGGTGGATTTCGAAGTTAAACGGTCAATTTGTTAAAG	556		
DB	161	TCAACGGAACACAGTTGAAGTTCATGAAGTTCATTCACACGTAAACGGAAAGAAAGTAAAG	220		
QY	557	TTTCTGCTGAACGCAACCAACATTCAGTCGGCTACTGATGGCGTAGAAATCGTTTC	616		
DB	221	TATTAGCTAACCGCAACCCAGAAGAAATTACCATGGGGCGAACTAGGGCGTAGATATCGTTTC	280		
QY	617	TTGAAGCAACTAGTTTCTTTTGCTAAAAAAGAGCTGCTGAAAAAACACATTACATGCTAACG	676		
DB	281	TAGAAATGACTGGTTTCTTCACATCTTAAATCAGCAGCTGAAAAACATTTAACAGCT--G	337		
QY	677	GTGCTAAAAAAGTTGTTATCACAGCTCCTGGTGGAAAAACGACGTTAAACAGTTGTTTCA	736		
DB	338	GTGCTAAACGCGTTGTTATCTCTGCTCCTGGCGAAAAACGATGTTCCAACAAATCGTTTATA	397		
QY	737	ACACTAACCCAGACATCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA	796		
DB	398	ACACAAACCCAGAAACATTAACCTGGTAAAGAAACAGTTATCTCAGGTGCTTCATGTACTA	457		
QY	797	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTCACGATGCAATTTGGTATCCAAAAAGGTC	856		
DB	458	CTAACTGTTTAGCTCCTATGGCTAAACATTTGAACGACAAAATTTGGTGTGTGTAAGGAT	517		
QY	857	TTATGACTACAAATCCACGCTTATACTGGTGACCAAAATGATCCTTGACGGACCAACCGTG	916		
DB	518	TAATGACAACTATCCACGCTTACACAGGTGACCAAAATGACTCTAGACGGACCTCATCTA	577		
QY	917	GTGGTGACCTTCGTGCTCGTGTGCTGCTGCTGCAAAACATTTGTTCCCTAACTCAACTGGTG	976		
DB	578	AAGGTGACTTCCGCGGTGCAACGCGTGTCTGCTGCAAAACATCGTTCCCTAACTCAACAGGTG	637		
QY	977	CTGCTAAAGCTATCGGTCTTTGTTATCCAGAAATGAATGGTAAACCTTGATGGTGCTGCAC	1036		
DB	638	CTGCTAAAGCTATCGGTTTAGTAATTCCTGAATTAACGGTAAATTAACGGAGCTGCTC	697		
QY	1037	AACGTGTTCTGTTCCAACTGGATCAGTAACCTGAGTTGGTTGTAACTCTTGATAAAAACG	1096		
DB	698	AACGTGTTCTGTACCAACAGGTTTCATTAAACAGAAATAGTAACAGATTCCTTGAAAAAAGAG	757		
QY	1097	TTTCTGTTGACGAATCAACGCTGCTATGAAAGCTGCTTCAAAACGACAGTTTCGGTTACA	1156		
DB	758	TAACTGTTGACGAATCAATGCAGCAATGAAGAAGCTTCAACAGAACTCTTACGGATACA	817		
QY	1157	CTGAAGATCCAAATGTTTCTTCAGATATCGTAGGCGGTGTCATACGGTTTCATTGTTTGACG	1216		
DB	818	ACACAGACGAATCGTTTCTTCAGATATCGTTGGTATGACTTACGGTTCAATTATTTGATG	877		
QY	1217	CAACTCAAACTAAAGTTATGGAAGTTGACGGATCAAAATGGTTAAAGTTGTATCATGGT	1276		
DB	878	CTACACAAACTAAAGTAATGACAGTTGGCGCAAAACAAATAGTTAAAAACTGTTGCTTGGT	937		
QY	1277	ATGACAAATGAAATGCTTACACTGCTCAACTTGTTCGTACACTTGAGTATTTTGCAA	1334		
DB	938	ACGACAAACGAATGTATACACTGCACAAATTAGTACGTAATTTAGAAATACCTTCGCTAA	995		

RESULT 13
US-09-134-000C-995
; Sequence 995, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134, 000C
; CURRENT FILING DATE: 1998-08-13

```

; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 995
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-995

```

	Query Match	35.5%;	Score 478;	DB 4;	Length 1041;
	Best Local Similarity	74.0%;	Pred. No. 7.6e-122;		
	Matches 620;	Conservative	0;	Mismatches 215;	Indels
QY	497	TCGATGGTACTGTTGAAGTTAAAGAAAGGTGGATTCGAAGTTAAACGGTCAATTTGTTAAAG	556		
DB	200	TCAACGGTACAGTTGAAGTTTACGGAAGTTCATTCACCGTTAAACGGCAAGAAATCAAAG	259		
QY	557	TTTCTGCTGAACGCCGAACCAACAGCAATTCAGCTGGGCTACTGATGGCGTAGAATCGTTC	616		
DB	260	TTTTAGCTAACCGTAAACCCTGAAGAAATTACCATGGGCGAATTAGGCGTAGACATCGTTT	319		
QY	617	TTGAAGCAACTAGTTTCTTTTGCTAAAAAGAAGCTGCTGAAAAACACTTACATGCTAACG	676		
DB	320	TAGAATGTAAGTGTCTTTTACTTCTAAAGAAGCTGCTGAAAAACATTTAACTGCT--G	376		
QY	677	GTGCTAAAAAAGTTGTTTATCACAGCTCCTGTGGAAAAACGACGTTAAAAACAGTTGTTTCA	736		
DB	377	GTGCAAAACGTTAGTTATTTTTCAGCTCCTGTGGTAAACGATGTACCAACAATCGTTTACA	436		
QY	737	ACACTAACCCAGCACATTCCTTGACGGTACTGAAACAGTTTATCTCAGGTGCTTCATGFACTA	796		
DB	437	ACACAAACCATGAACATTAACCTGGAGAGAACTGTAAATTTTCAGGCGCTTCTTTGFACTA	496		
QY	797	CAAACTGTTTAGTCTCTATGGCTAAAGCTCTTCACGATGCATTTGGTATCCAAAAAGGTC	856		
DB	497	CAAACTGCTTAGCTCCAATGGCTAAAGCTTTACATGACAACTTTGGTGTGTTGAAGGTT	556		
QY	857	TTATGACTACAATCCACGCTTATCTGTTGACCAAAATGATCCTTGACGGACCAACCGTG	916		
DB	557	TAATGACAACTATCCACGCTTACACAGGTGACCAAAATGACATTAGACGGACCAACATCCTA	616		
QY	917	GTGGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	976		
DB	617	AAGCGACCTTCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	676		
QY	977	CTGCTAAAGCTATCGGTCTTGTATTATCCAGAAATGAATGGTAAACCTTGATGGTGTGCAC	1036		
DB	677	CTGCTAAAGCAATCGGCTTAGTAATCCAGAAATGAACGGTAAATAGATGGCGCTGCTC	736		
QY	1037	AACGTGTTCCGTTCCTGTTCCAACTGGATCAGTAACAGTGGTGGTTGTTGTAACCTCTTGATAAAAACG	1096		
DB	737	AACGTGTTCCGTGCTAGCAACTGGTTTCACTGAATAGTTACTGTATTTAGACAAAGAAG	796		
QY	1097	TTTCTGTTGACGAAATCAACGCTGCTATGAAGCTGCTTCAAAACGACAGTTTCGGTTACA	1156		
DB	797	TTACTGTTGATGAAGTAAATGCAGTAATGGAAAAAGCTGCTAACGAATCTTATGGTTATA	856		
QY	1157	CTGAAGATCCAATTGTTTCTTCAGATATCGTAGGCGTGTCTACACGGTTCATGTTTGACG	1216		
DB	857	ACACAGACGAAATCGTTTCTTCTGATATCGTAGGTATGACTTACGGTTTCATTTATTCGATG	916		
QY	1217	CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATGGTTAAAGTTGTATCATGGT	1276		
DB	917	CAACTCAAACTAAAGTTGATGACAGTTGGCGACAAACAATAGTTTAAACCTGTTGCTTGGT	976		
QY	1277	ATGACAAATGAATGTCTTACACTGCTCAACTTGTTCGTACACTTGAGTATTTTGCAA	1334		
DB	977	ATGACAAACGAAATGTATACACTGCTCAATTAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1034		

RESULT 14
US-09-134-001C-2676

	Query Match	31.7%;	Score 426.8;	DB 4;	Length 1047;
	Best Local Similarity	70.4%;	Pred. No. 9.4e-108;		
	Matches 601;	Conservative 0;	Mismatches 247;	Indels 6;	Gaps 2;
QY	497	TCGATGGTACTGTTGAAGTTAAAGAAAGGTGGATTTCGAAGTTAAACGGTCAATTTGTTAAAG	556		
Db	197	TCAC TGGAGAAAGTTGAAGTTATCGAAGGTGGATTCGGTGTGAACGGTAAAGAAATTAAAT	256		
QY	557	TTTCTGCTGAACGCGGAACCCAGCAAAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC	616		
Db	257	CATTTCGATGAACCCAGATCGTGGTAAATTACCATGGGGCGATTTAGATATCGACCGTAGTAT	316		
QY	617	TTGAAGCAACTAGTTTCTTTTGCTAAAAAAGAAAGCTGCTGAATAAACACATTACATGCTAAACG	676		
Db	317	TAGAATGTACTGGTTTCTATACTGATAAAGAAAAAGCAAGCTCACATCGATGCG---AG	373		
QY	677	GTGCTAAAAAAGTTGTTATCACAGCTCCTGGTGGAAACGACGTTAAACAGTTGTTTTCA	736		
Db	374	GTGCTAAAAAAGTATTAATCTCAGCTCCAGCTAAAGGTGATGTAAAAACAATCGTATTCA	433		
QY	737	ACACTAACCCAGCAATCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA	796		
Db	434	ACACTAACCCATGATACATTAGATGGTTTCAGAAAAACAGTTGTTTCAGGTGCTTCTTGTACTA	493		
QY	797	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTTCACGATGCAATTTGGTATCCAAAAAGGTC	856		
Db	494	CTAACTCATTAGCACCAAGTTGCCAAAAGTTTAAAGTGACGAATTCGGTTTAGTTGAAGGTT	553		
QY	857	TTATGACTACAATCCACGCTTATACTGGTGACCAATGATCCTTGACGGACCACACCGTG	916		
Db	554	TCATGACTACAATTCACGCTTACACTGGTGACCAAAATACACAAGACGCACCTCACAGAA	613		
QY	917	GTGGTGACCTTCGTCGTGCTCGTGGTGCTGCAAAACATTTCTCCTAACTCAACTGGTG	976		
Db	614	AAGGTGACAAACGTCGTGCACGTGCAGCAGCTGAAAATATTATCCCTAACTCAACAGGTG	673		
QY	977	CTGCTAAAGCTATCGGTCTGTTATCCCAGAAATTGAATGGTAAACTTTGATGGTGTGCAC	1036		
Db	674	CTGCTAAAGCTATCGGTAAAAGTTATTCCAGAAATCGATGGTAAATTAGACGGTGGAGCAC	733		
QY	1037	AACGTGTTCCCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTTAACTCTTGATA---AAA	1093		
Db	734	AACGTGTTCCAGTTGCTACTGTTCTTTTAACTGAATTAAGTGTAGTATTAGACAAACAAG	793		
QY	1094	ACGTTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAACGACAGTTTCGGTT	1153		
Db	794	ATGTAACGTGTTGACCAAGTTAAACAGTGCTATGAACAAGCTTCTGACGAATCAITTCGGTT	853		
QY	1154	ACACTGAAGATCCAAATTGTTTCTTCAGATATCGTAGGCGTGTATACCGTTTCATTGTTTG	1213		
Db	854	ACACTGAAGACGAAAATCGTATCTTCTGATATTGTTGGTATGACTTACCGTTTCATTATTG	913		
QY	1214	ACGCAACTCAAACTAAAGTTATGGAAGTTGACGGATCAAAATGGTTAAAGTTGTATCAT	1273		

[illegible]

QY	617	TTGAAGCAACTAGTTCTTTTGCTAAAAAAGAAGCTGCTGAAAAACACTTACATGCTAACG	676
Db	1472	TAGAATGTACTGGTTTCTACACTGATAAAGATAAAGCACAAAGCTCATATTGAAGCA---	1528
QY	677	GTGCTAAAAAAGTTGTTATCACAGCTCCTGGTGGAAACGACGTTAAAAACAGTTGTTTCA	736
Db	1529	GGCTAAAAAAGTATTAACTCTCAGCACCACTACTGGTGACTTAAAAACAATCGTATTCA	1588
QY	737	ACACTAACCCACGACATTCTTGACGGTACTGAAAAACAGTTATCTCAGGTGCTTCATGTACTA	796
Db	1589	ACACTAACCCACCAAGAGTTAGACGGTTCTGAAAAACAGTTGTTTCAGGTGCTTCATGTACTA	1648
QY	797	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTCACGATGCATTTGGTATCCAAAAAGGTC	856
Db	1649	CAAACTCATTTAGCACCAAGTTGCTAAAGTTTAAACGATGACTTTGGTTTAGTTGAAGGTT	1708
QY	857	TTATGACTACAATCCACGCTTATCTGCTGACCAAAATGATCCTTGACGGACCAACCGTG	916
Db	1709	TAATGACTACAATTCACGCTTACACAGGTGATCAAAATACACAAGACGCACCTCACAGAA	1768
QY	917	GTGGTGACCTTCGTGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	976
Db	1769	AAGGTGACAAAACGTCGTGCTCGTGCGCGGAGAAACATCATCCCTAACTCAACAGGTG	1828
QY	977	CTGCTAAAGCTATCGGTCTTGTATCCACAGAAATGAATGGTAACTGATGGTCTGCAC	1036
Db	1829	CTGCTAAAGCTATCGGTAAAGTTATTCCTGAAATCGATGGTAAATAGATGGTGTGCAC	1888
QY	1037	AACGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1093
Db	1889	AACGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1948
QY	1094	ACGTTTCTGTTGACGAAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1153
Db	1949	ACGTTAACAGTTGACGAAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2008
QY	1154	ACACTGAAGATCCAAATGTTTCTTCAGATATCGTAGGCGTGTATACGGTTTCATTGTTTG	1213
Db	2009	ACACTGAAGACGAAATCGTTTCTTCAGACGTTGTAGGTATGACTTACGGTTTCATTATTCG	2068
QY	1214	ACGCAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGTATCAT	1273
Db	2069	ACGCTACACAAACTCGTGAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2128
QY	1274	GGTATGACAAATGAATGTCTTACACTGCTCAACTTGTTCGTACACTTGAGTATTTTGCAA	1333
Db	2129	GGTATGATAACGAAATGTCTATATACTGCACAAATAGTTTCGTACATTAGCATACTTAGCTG	2188
QY	1334	AAATCGCTAAATAA	1347
Db	2189	AACTTCTAAATAA	2202

Search completed: March 31, 2004, 03:10:34
Job time : 150 secs

	Query Match	100.0%;	Score 1347;	DB 9;	Length 1347;
	Best Local Similarity	100.0%;	Pred. No. 2.6e-308;		
	Matches 1347;	Conservative	0;	Mismatches	0;
				Indels	Gaps
				0;	0;
Qy	1	ATCAAAAAAAAAA	TACAGGGATTATTTATTTGCTTCTTGCAGTCATTATTCTGTCTGCATGC	60	
Db	1	ATCAAAAAAAAAA	TACAGGGATTATTTATTTGCTTCTTGCAGTCATTATTCTGTCTGCATGC	60.	
Qy	61	CAGSCAAACTACGGATCCGGTATCGT	TAGTTAAAGTTGGTATTAAACGGTTTCGGTTCGTATC	120	
Db	61	CAGSCAAACTACGGATCCGGTATCGT	TAGTTAAAGTTGGTATTAAACGGTTTCGGTTCGTATC	120	

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1347	100.0	1347	9	US-09-878-766A-21	Sequence 21, Appl
2	815.8	60.6	1011	9	US-09-878-766A-11	Sequence 11, Appl
3	815.8	60.6	1011	10	US-09-878-781-3	Sequence 3, Appli
4	815.8	60.6	1011	14	US-10-134-297-3	Sequence 3, Appli
5	790.2	58.7	961	10	US-09-878-781-13	Sequence 13, Appl
6	780.6	58.0	1011	12	US-10-282-122A-38195	Sequence 38195, A
7	778.2	57.8	1010	10	US-09-878-781-15	Sequence 15, Appl
8	702.2	52.1	1011	9	US-09-878-766A-13	Sequence 13, Appl
9	702.2	52.1	1011	10	US-09-878-781-5	Sequence 5, Appli
10	702.2	52.1	1011	14	US-10-134-297-5	Sequence 5, Appli
11	683	50.7	1080	9	US-09-815-242-9495	Sequence 9495, Ap
12	678.2	50.3	1011	9	US-09-878-766A-19	Sequence 19, Appl
13	678.2	50.3	1011	10	US-09-878-781-11	Sequence 11, Appl
14	678.2	50.3	1011	14	US-10-134-297-11	Sequence 11, Appl
15	678.2	50.3	1080	9	US-09-815-242-9071	Sequence 9071, Ap

QY	121	GGACGCTTTGCAATCCGTCGTATTCAAATGTTGAAGGTTGTTGAAGTAACTCGTATCAAC	180
DB	121		
QY	181	GACCTTACAGATCCAAACATGCTTGCCACACTTGTGAAATACGATACAACCTCAAGGACGT	240
DB	181		
QY	241	TTTGAACGGAACCTGTTGAAGTTAAAGAAAGGTGGATTGGAAGTAAACGGAAAACTTCATCAAA	300
DB	241		
QY	301	GTTTTCTGCTGAACGTGATCCAGAAAAACATCGACTGGGCAACTGACGGTGTGAAATCGTT	360
DB	301		
QY	361	CTGGAAGCACTCGAGGGTACTGTAGAAAGTTAAAGATGGTGGATTGACGTTAAACGGAATA	420
DB	361		
QY	421	TTCAATTAAGTTTCTGCTGAAAAAGATCCAGAACAAATTGACTGGSCAACTGACGGTGT	480
DB	421		
QY	481	GAAATCGTTCTTGAATTCGATGGTACTGTTGAAAGTTAAAGAAAGGTGGATTGGAAGTTAA	540
DB	481		
QY	541	GGTCAATTTGTTAAAGTTTCTGCTGAACGGCAACGAGCAAAACATTGACTGGGCTACTGAT	600
DB	541		
QY	601	GGCGTAGAAATCGTTTCTTGAAGCAACTAGTTTCTTTGCTAAAAAGAAAGCTGCTGAAAA	660
DB	601		
QY	661	CACCTACATGCTAACCGTGCTAAAAAGTTGTTATCACAGCTCCTGGTGGAAACGACGTT	720
DB	661		
QY	721	AAACACAGTTGTTTTCAACACATAACCCAGACATTCTTGACGGTACTGAAACAGTTATCTCA	780
DB	721		
QY	781	GGTGCTTCATGTACTACAAAACCTGTTTAGCTCCTATGGCTAAAGCTCTTCAAGATGCATTT	840
DB	781		
QY	841	GGTATCCAAAAAGGTTCTATGACTACAATCCACGCTTATCTGGTGACCAAAATGATCCTT	900
DB	841		
QY	901	GACGGACCAACACCGTGGTGACCTTCGTGCTGCTGCTGGTGCTGCAAAACATTTGTT	960
DB	901		
QY	961	CCTAACTCAACTGGTGTCTAAAGCTATCGGTCTTGTATCCAGAAATGAAATGGTAAA	1020
DB	961		
QY	1021	CTTGATGGTGCTGCACAAACGTGTTCCCTGTTCCAACTGGATCAGTAACTGAGTTGTTGTA	1080
DB	1021		
QY	1081	ACTCTTGATAAAAAACGTTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAAC	1140
DB	1081		
QY	1141	GACAGTTTCGGTTACACTGAAGATCCAAATTTGTTTCTTCAGATATCGTAGGCGTGTATAC	1200
DB	1141		
QY	1201	GGTTCATGTTTGAACGCAACTCAAACCTAAAGTTATGGAAGTTGACGGATCACAATTTGGTT	1260

RESULT 2

```

US-09-878-766A-11
; Sequence 11, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY C
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus dysgalactiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1011)
;
US-09-878-766A-11

```

Query Match	60.6%;	Score 815.8;	DB 9;	Length 1011;
Best Local Similarity	97.4%;	Pred. No. 9.7e-183;		
Matches 829; Conservative	0;	Mismatches 22;	Indels 0;	Gaps 0;

[illegible]

QY 1337 TCGCTAAATAA 1347
| | | | |
Db 951 TTGCTAAATAA 961

RESULT 6

US-10-282-122A-38195
; Sequence 38195, Application US/10282122A
; Publication No. US20040029129A1

GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 38195

; LENGTH: 1011

; TYPE: DNA

; ORGANISM: Streptococcus pyogenes

US-10-282-122A-38195

Query Match 58.0%; Score 780.6; DB 12; Length 1011;

Best Local Similarity 94.8%; Pred. No. 2e-174;

Matches 807; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 497 TCGATGGTACTGTTGAAGTTAAAGAGGTGGATTCCGAGTTAAACGGTCAATTGTTAAAG 556

| | | | |

Db 161 TTGATGGAACAGTTTGAAGTTAAAGAGGTGGATTTCGAGTTAAACGGAACTTCATCAAG 220

| | | | |

QY 557 TTTCTGCTGAACCGCAACCCAGCAACATTGACTGGGCTACTGATGGCGTAGAATCGTTC 616

| | | | |

Db 221 TTTCTGCTGAACGTGATCCAGAAAACATCGACTGGGCAACTGATGGGTTGAAATCGTTC 280

| | | | |

QY 617 TTGAAGCAACTAGTTTCTTTGCTAAAAAAGAGCTGCTGAAAAACACATTACATGCTAACG 676

| | | | |

Db 281 TTGAAGCAACTGGTTTCTTTGCTAAAAAAGAGCTGCTGAAAAACACATTACATGCTAACG 340

| | | | |

QY 677 GTGCTAAAAAAGTTGTTATCACAGCTCCTGGTGGAAACGACGTTAAAAACAGTTGTTTCA 736

| | | | |

Db 341 GTGCTAAAAAAGTTGTTATCACAGCTCCTGGTGGAAACGATGTTAAAAACAGTTGTTTCA 400

QY 737 AACTAACCCAGACATTTCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA 796

| | | | |

Db 401 AACTAACCCAGACATTTCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA 460

| | | | |

QY 797 CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTCCAGATGCTATGGTATCCAAAAAGTTC 856

| | | | |

Db 461 CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTCCAGATGCTATGGTATCCAAAAAGTTC 520

| | | | |

QY 857 TTATGACTACAAATCCACGCTTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916

| | | | |

Db 521 TTATGACTACAAATCCACGCTTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 580

| | | | |

QY 917 GTGGTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 976

| | | | |

Db 581 GTGGTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 640

| | | | |

QY 977 CTGCTAAAGCTATCGGTCTTGTGTTATCCAGAAATTGAATGGTAAACTTGTGCTGCAC 1036

| | | | |

Db 641 CTGCTAAAGCTATCGGTCTTGTGTTATCCAGAAATTGAATGGTAAACTTGTGCTGCAC 700

| | | | |

QY 1037 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTTAACTCTTGATAAAAACG 1096

| | | | |

Db 701 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTTAACTCTTGATAAAAACG 760

| | | | |

QY 1097 TTTCTGTTGACGAAATCAACGGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGGTTACA 1156

| | | | |

Db 761 TTTCTGTTGACGAAATCAACGGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGGTTACA 820

| | | | |

QY 1157 CTGAAGATCCAATTTGTTCTTCCAGATATCGTAGGGCTGTATACGGTTTATGTTGACG 1216

| | | | |

Db 821 CTGAAGATCCAATTTGTTCTTCCAGATATCGTAGGGCTGTATACGGTTTATGTTGACG 880

| | | | |

QY 1217 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGTTAAAGTTGATCATGGT 1276

| | | | |

Db 881 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGTTAAAGTTGATCATGGT 940

| | | | |

QY 1277 ATGACAAATGAAATGTCTTACACTGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTG 1336

| | | | |

Db 941 ATGACAAAGAAATGTCTTACACTGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1000

| | | | |

QY 1337 TCGCTAAATAA 1347

| | | | |

Db 1001 TTGCTAAATAA 1011

RESULT 7

US-09-878-781-15

; Sequence 15, Application US/09878781

; Publication No. US20030082781A1

; GENERAL INFORMATION:

; APPLICANT: Bolton, Alexandra J.

; APPLICANT: Perez-Casal, Jose

; APPLICANT: Fontaine, Michael

; APPLICANT: Potter, Andrew A.

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST

; TITLE OF INVENTION: STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0055

; CURRENT APPLICATION NUMBER: US/09/878,781

; CURRENT FILING DATE: 2002-09-10

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 15

; LENGTH: 1010

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SeqGapC

US-09-878-781-15

Query Match

Best Local Similarity 57.8%; Score 778.2; DB 10; Length 1010;

Pred. No. 7.5e-174;

RESULT 12
US-09-878-766A-19
; Sequence 19, Application US/09878766A
; Patent No. US2002004928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus iniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1011)
US-09-878-766A-19

Query Match 50.3%; Score 678.2; DB 9; Length 1011;
Best Local Similarity 87.3%; Pred. No. 3.4e-150;
Matches 743; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 497 TCGATGCTACTGTTGAAGTTAAAGAGGTGGATTGCGAAGTTAACGGTCAATTTGTTAAAG 556
Db 161 TTGACGGTACAGTTGAAGTTAAAGATGGTGGATTGCGAAGTTAACGGAGCTTTGTTAAAG 220

QY 557 TTTCTGCTGAACGCGAACCGAACCAGCAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC 616
Db 221 TTTCTGCGAAGCGGACCGAACCGAACCAGCAACATTGACTGGGCTACTGATGGTGTAGACATCGTTC 280

QY 617 TTGAAGCAACTAGTTTCTTTGCTTAAAGAAAGAGCTGCTGTAAGAAACACTTACATGCTAACG 676
Db 281 TTGAAGCAACAGGTTTCTTCGCTTCTAAAGCAGCTGCTGTAAGCAACACATTCACGCTAACG 340

QY 677 GTGCTAAAAAGTTGTTATCACAGCTCCTGGTGAAGAACGCTTAAACAGTTGTTTCA 736
Db 341 GTGCGAAAAAGTTGTTATCACAGCTCCTGGTGAAGAAATGAGTTAAACAGTTGTTTACA 400

QY 737 ACACCTAACCGACACATTCTTGACGGTACTGAAACAGTTTATCTCAGGTGCTTCATGTAATA 796
Db 401 ACACCTAACCGACATATTTCTTGATGGAACAGTTTATCTCAGGTGCTTCATGTAATA 460

QY 797 CAAACTGTTAGCTCCTATGGCTTAAAGCTTTCACGATGCTTAAAGAAAGCTTAAAGAAAGCTC 856
Db 461 CAAACTGTTAGCTCCTATGGCTTAAAGCTTTCACGATGCTTAAAGAAAGCTTAAAGAAAGCTC 520

QY 857 TTATGACTACAAATCCACGCTTATCTGTTGACCAAAATGATCTTACGGGACCAACCGCTG 916
Db 521 TAATGACTACTATCCATGTTTACACTGGTGACCAAAATGTTTCTTGACGGACCAACCGCTG 580

QY 917 GTGGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 581 GTGGTGATCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640

QY 977 CTGCTAAAGCTATCGGTCTGTTATCCAGAAATGAAATGGTAAAGCTTATGATGGTGTGCTGAC 1036
Db 641 CTGCTAAAGCAATCGGTCTGTTATCCAGAAATGAAATGGTAAAGCTTATGATGGTGTGCTGAC 700

QY 1037 AACGTGTTCTGTTCCAACTGGATCAGTAAGTGGTGGTTGTTGTAAGCTTCTGATAAAAGCT 1096
Db 701 AACGTGTTCTGTTCCAACTGGATCAGTAAGTGGTGGTTGTTGTAAGCTTCTGATAAAAGCT 760

QY 1097 TTTCTGTTGAGCAAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGGTTTACA 1156
Db 761 CTTGAGTAGAAGAAATCAATGCAGCTATGAAAGCAGCAGCTTCAACGATTCATACGGTTTACA 820

QY 1157 CTGAAGATCCAAATGTTTCTTTCAGATATCGTAGGGTGTCTATACGGTTCATTTGTTGACG 1216
Db 821 CTGAAGATGCTATCGTATCATCATGATATCGTAGGTATTTCTTACGGTTCATTTGTTGATG 880

QY 1217 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCAACAATGGTTAAAGTTGTATCATGGT 1276
Db 881 CTACTCAAACTAAAGTACAAACTGTTGATGGAATCAATGGTTAAAGTTGTTTCATGGT 940

QY 1277 ATGACAAATGAAATGTCTTACACTGCTCAACTTGTTCGTACACTTGTGAGTATTTTGCAAAA 1336
Db 941 ATGACAAATGAAATGTCTTACACTGCTCAACTTGTTCGTACTCTTGAGTACTTTTGCAAAA 1000

QY 1337 TCGCTAAATAA 1347
Db 1001 TCGCTAAATAA 1011

RESULT 13
US-09-878-781-11
; Sequence 11, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus iniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1011)
US-09-878-781-11

Query Match 50.3%; Score 678.2; DB 10; Length 1011;
Best Local Similarity 87.3%; Pred. No. 3.4e-150;
Matches 743; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 497 TCGATGCTACTGTTGAAGTTAAAGAGGTGGATTGCGAAGTTAACGGTCAATTTGTTAAAG 556
Db 161 TTGACGGTACAGTTGAAGTTAAAGATGGTGGATTGCGAAGTTAACGGAGCTTTGTTAAAG 220

QY 557 TTTCTGCTGAACGCGAACCGAACCAGCAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC 616
Db 221 TTTCTGCGAAGCGGACCGAACCAGCAACATTGACTGGGCTACTGATGGTGTAGACATCGTTC 280

QY 617 TTGAAGCAACTAGTTTCTTTGCTTAAAGAAAGAGCTGCTGAAAAACACTTACATGCTAACG 676
Db 281 TTGAAGCAACAGGTTTCTTCGCTTCTTAAAGCAGCTGCTGAAACACACATTCACGCTAACG 340

QY 677 GTGCTAAAAAGTTGTTATCACAGCTCCTGGTGAAGAACGCTTAAAGAAAGCTTAAAGAAAGCT 736
Db 341 GTGCGAAAAAGTTGTTATCACAGCTCCTGGTGAAGAAATGAGTTAAACAGTTGTTTACA 400

QY 737 ACACCTAACCGACACATTCTTGACGGTACTGAAACAGTTTATCTCAGGTGCTTCATGTAATA 796
Db 401 ACACCTAACCGACATATTTCTTGATGGAACAGTTTATCTCAGGTGCTTCATGTAATA 460

QY 797 CAAACTGTTAGCTCCTATGGCTTAAAGCTTTCACGATGCTTAAAGAAAGCTTAAAGAAAGCTC 856
Db 461 CAAACTGTTAGCTCCTATGGCTTAAAGCTTTCACGATGCTTAAAGAAAGCTTAAAGAAAGCTC 520

QY 857 TTATGACTACAAATCCACGCTTATCTGTTGACCAAAATGATCTTACGGGACCAACCGCTG 916
Db 521 TAATGACTACTATCCATGTTTACACTGGTGACCAAAATGTTTCTTGACGGACCAACCGCTG 580

QY	497	TCGATGGTACTGTTGAAGTTAAAGAAAGGTGGATT	CGAAGTTAAACGGTCAATTTGTTAAAG	556
DB	233	TCGACGGTACTGTTGAAGTTAAAGAAAGGTGGATT	TGAAGTTAAACGGTAAATTCATCAAAG	292
QY	557	TTTCTGCTGAACGCGAACCAGCAAAACATTGACT	CTGGGTACTGATGGCGTAGAAATCGTTTC	616
DB	293	TTTCTGCTGAACGCTGATCCAGAAACAAATCGAC	CTGGGTACTGACGGTGTAGAAATCGTTTC	352
QY	617	TTGAAGCAACTAGTTTTCTTTTGCTAAAAAAGAA	AGCTGCTGAGAAAAACACATTACATGCTAAACG	676
DB	353	TTGAAGCTACTGGTTTCTTTTGCTAAGAAAGAA	GCAGCTGAAAAACACCTTAAAGGT---G	409
QY	677	GTGCTAAAAAAGTTGTTATCACAGCTCCTGGT	GGAACGACGTTAAAAACAGTTGTTTCA	736
DB	410	GAGCTAAAAAAGTTGTTATCACTGCTCCTGGT	GGAACGACGTTAAAAACAGTTGTTTCA	469
QY	737	ACACTAACCCAGCACATTCCTTGACGGTACTGAA	ACAGTTATCTCAGGTGCTTCATGTACTA	796
DB	470	ACACTAACCCAGCAGCTTCTTGACGGTACTGAA	ACAGTTATCTCAGGTGCTTCATGTACTA	529
QY	797	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTC	CAGGATGCATTTGGTATCCAAAAAGGTC	856
DB	530	CAAACTGCTTGGCTCCAAATGGCTAAAGCTCTT	CAAGCAACACTTTGGTGTGTTGAAGGAT	589
QY	857	TTATGACTACAATCCACGCTTATCTGGTGACCA	AAATGATCCTTGACGGACCACACCGTG	916
DB	590	TGATGACTACTATCCACGCTTACACTGGTGAC	CAATGATCCTTGACGGACCACACCGTG	649
QY	917	GTGGTGACCTTCGTCGTGCTGCTGGTGCTGCA	AAACATTTGTTCTTAACCTCAACTGGTG	976
DB	650	GTGGTGACCTTCGCCGCTGCTCGCGCTGGTGCT	GCAAAACATCGTTCTTAACCTCAACTGGTG	709
QY	977	CTGCTAAAGCTATCGGTCCTTGTATTATCCCAG	AATTGGAATGGTAAACCTTGATGGTGTGCAC	1036
DB	710	CTGCAAAAGCTATCGGTCCTTGTAAATCCCAG	AATTGGAATGGTAAACCTTGATGGTGTGCAC	769
QY	1037	AACGTGTTCCCTGTTCCAACTGGATCAGTAAC	TGAGTTGGTTGTAACCTTGTATAAAAACG	1096
DB	770	AACGCGTTCCAACTCCAACTGGATCAGTTACT	GAATGGTAGCAGTTCTTTGAAAAAGAACG	829
QY	1097	TTTCTGTTGACGAAATCAACGCTGCTATGAA	AGCTGCTTCAAAACGACAGTTTCGGTTACA	1156
DB	830	TTACTGTTGATGAAGTGAAGCAGCAGCTATGA	AAAGCAGCTTCAAAACGAAATCATACGGTTACA	889
QY	1157	CTGAAGATCCAATTTGTTTCTTTCAGATATCG	TAGCGGTGTCATACGGTTTGTGACG	1216
DB	890	CAGAAGATCCAATCGATCTTTCAGATATCGT	AGGTATGTTCTTACGGTTCTCATGTTTGACG	949

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 22:21:00 ; Search time 3333 Seconds
(without alignments)

12068.514 Million cell updates/sec

Title: US-10-650-369-21

Perfect score: 1347

Sequence: 1 atgaaaaaaataacagggat.....ttgcaaaaatcgctaaataa 1347

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	471.8	35.0	1127	28	BH770540 LLMGtag30
2	195.6	14.5	990	14	CB686102 Bn01b 02m
3	190.6	14.1	944	28	AF075927
4	189	14.0	973	14	CK261340 EST707418

C	5	187.6	13.9	892	28	AZ536323	AZ536323	ENTCQ33TR
	6	185.4	13.8	772	12	BI433946	BI433946	EST536707
	7	181.6	13.5	955	14	CK276937	CK276937	EST723015
	8	179.2	13.3	891	14	CK272330	CK272330	EST718408
	9	178.2	13.2	973	14	CK272824	CK272824	EST718902
	10	177.8	13.2	641	9	AI779267	AI779267	EST260146
	11	177.8	13.2	783	12	BI422981	BI422981	EST533647
	12	175.6	13.0	530	10	AW267163	AW267163	YA61 gast
	13	175.4	13.0	626	9	AI782605	AI782605	EST263484
	14	174	12.9	771	12	BQ045826	BQ045826	EST594944
	15	172.8	12.8	698	14	CA835550	CA835550	MCS038C08
	16	172.2	12.8	982	28	BH158712	BH158712	ENTRK84TR
	17	171.8	12.8	764	12	BC590989	BC590989	EST498831
	18	171.6	12.7	598	9	AW096431	AW096431	EST289611
	19	171.4	12.7	915	28	AZ692190	AZ692190	ENTMR86TF
	20	170.8	12.7	679	14	CA836228	CA836228	MCU004H12
	21	170.8	12.7	684	14	CA836289	CA836289	MCU005F09
	22	170.8	12.7	693	14	CA835984	CA835984	MCU002C12
	23	169.6	12.6	701	12	BJ400991	BJ400991	BJ400991
	24	169.6	12.6	705	12	BJ398055	BJ398055	BJ398055
	25	169.6	12.6	721	12	BJ374157	BJ374157	BJ374157
	26	169.6	12.6	722	12	BJ398333	BJ398333	BJ398333
	27	169.6	12.6	727	9	AU039919	AU039919	AU039919
	28	169.4	12.6	714	14	CA836856	CA836856	MCU011G04
	29	169.2	12.6	731	10	AW775569	AW775569	EST334634
	30	169.2	12.6	866	13	BQ704899	BQ704899	Bn01_01a2
	31	168.6	12.5	739	12	BJ384060	BJ384060	BJ384060
	32	168.2	12.5	724	9	AU060734	AU060734	AU060734
	33	168	12.5	575	9	AW093693	AW093693	EST286873
	34	168	12.5	720	12	BJ399702	BJ399702	BJ399702
	35	167.2	12.4	712	12	BM300046	BM300046	MCR054E07
	36	166.8	12.4	658	12	BI945772	BI945772	SR92a02.Y
	37	166.8	12.4	797	12	BM411386	BM411386	EST585713
	38	166.6	12.4	679	12	BG452149	BG452149	NF077F10L
	39	166.4	12.4	589	9	AW096506	AW096506	EST289686
	40	165.6	12.3	729	12	BJ430302	BJ430302	BJ430302
	41	165.6	12.3	784	10	BE130677	BE130677	L48-879T3
	42	165.4	12.3	589	12	BG452151	BG452151	NF077F11L
	43	165.2	12.3	694	12	BG268898	BG268898	L0-2874T3
	44	164.4	12.2	876	14	CK260911	CK260911	EST706989
	45	164.2	12.2	703	9	AU062136	AU062136	AU062136

ALIGNMENTS

RESULT 1
BH770540 1127 bp DNA linear GSS 01-MAY-2002

LOCUS LLMGtag303 MG1363 Random Sequence Tag Library Lactococcus lactis

DEFINITION subsp. cremoris genomic, genomic survey sequence.

ACCESSION BH770540 GI:20373497

VERSION BH770540

KEYWORDS GSS.

SOURCE Lactococcus lactis subsp. cremoris

ORGANISM Lactococcus lactis subsp. cremoris

REFERENCE 1 (bases 1 to 1127)

AUTHORS Bolotin,A., Ehrlich,S.D. and Sorokin,A.

TITLE Studies of genomes of dairy bacteria Lactococcus lactis

JOURNAL Sci. Aliments (2002) In press

COMMENT Contact: Sorokin A
Genetique Microbienne
INRA
CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21
Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is gapA (99%)
Class: shotgun
High quality sequence start: 30
High quality sequence stop: 1099.

Db 369 CCATGAAGATACAAATCATCAGCAACGCCTCTGTACTACTAACTGCTCGTCCATTGCT 428

QY 819 TAAAGCTCTTCACGATGCAATTTGGTATCCAAAAGGTCTTATGACTACAATCCACGCTTA 878

Db 429 CAAGGTTCTTGACAGAAATTCGGGATCATCAAGGGGCAATATGACACCACTCACTCATA 488

QY 879 TACTGGTGACCAATGATCCTTTGACGGACCAACCGGTGGTGGTGACCTTCGTGCTCG 938

Db 489 CACTGGTGACCAAGGCTGTTGGATGCAAGCCACC-----GTGATCTAAGGAGAGCAAG 542

QY 939 TGCTGGTGCTGCAACATGTTCTCTAACTCAACTGGTGTGCTGTCTAAAGCTATCGGTTCT 998

Db 543 AGCAGCTGCTTTAAACATCGTTCCAACTCAACAGGAGCAGCTAAAGCCGTGCTCTTGT 602

QY 999 TATCCAGAAATGATGTTAAACTTGTATGCTGCTGCACAAACGTTGTTCTGTTCCAACTGG 1058

Db 603 CCTCCCTAACCTCAAGGAAAGCTCAACGSCATCGCATTCGTTGCCAACTCCCAACGT 662

QY 1059 ATCAGTAACTGAGTTGTTGTAACCTCTTGATAAAAACGTTTCTGTTGACGAAATCAACGC 1118

Db 663 CTCGGTGTGACTTAGTTGTGCAAGTCTCCCAAGAGACTTTTGTCTGAAGAAGTCAACGC 722

QY 1119 TGC---TATGAAAGCTGTTCAAACGACAGTTTCGGTTACACTGAAGATCCAATGTTTC 1175

Db 723 TCGGTTACGGGATGCAGTCCGAACGAGCTTAAGGTATACTCGATGTCTCGGATGAGCC 782

QY 1176 TTCAGATATCGTAGGCTGCATACGGTTCAATTTGTTGACGCAACTCAAACTAAAGTTAT 1235

Db 783 TCTTGCTCTGTCGATTCAGGTGCTCTGATGTTGTCATCCACCATTTGATTTCTCTCAC 842

QY 1236 GGAAGTTGACGGATCAACAATTTGTTAAAGTTGATCATGTTGATGACAAATGATCTTA 1295

Db 843 AATGTTATGGGAGATGATATGTTAAAGTATGCTTGGTATGATATGATATGATGGGCTA 902

QY 1296 CACTGCTCAACTTGTTCGTACACTTGAGTATTTTGCAAAATCGCTAAATAA 1347

Db 903 CTCTCAGAGAGTTGTTGATTGGCTGACATTTGTTGCCAATAACTGGAAAGTGA 954

RESULT 3

AF075927

LOCUS

DEFINITION

AF075927

AF075927.1 GI:3320797

GSS.

Salmonella typhimurium

Salmonella typhimurium

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.

1 (bases 1 to 944)

Wong,R.M.-Y., Wong,K.K., Benson,N.R. and McClelland,M.

Sample sequencing of a Salmonella typhimurium LT2 lambda library: comparison to the Escherichia coli K12 genome

FEMS Microbiol. Lett. 173 (2), 411-423 (1999)

99243757

10227170

Contact: McClelland M

Molecular Biology

Sidney Kimmel Cancer Center

3099 Science Park Road, San Diego, CA 92121, USA

Email: mclelland@lifsci.sdsu.edu

Class: shotgun.

Location/Qualifiers

1. .944

/organism="Salmonella typhimurium"

/mol_type="genomic DNA"

/strain="LT2"

/db_xref="taxon:602"

/clone="B159-T7"

/clone_lib="Salmonella typhimurium LT2, Lambda DASH II"

FEATURES

source

/note="Vector: Lambda DASH II; sequenced using Li-Cor sequencer"

ORIGIN

Query Match 14.1%; Score 190.6; DB 28; Length 944;

Best Local Similarity 57.7%; Pred. No. 4.8e-38;

Matches 469; Conservative 0; Mismatches 314; Indels 30; Gaps 6;

QY 497 TCGATGGTACTGTTGAAGTTAAAGAGGTGATTCGAAAGTTAACGGTCAATTTGTTAAAG 556

Db 65 TCGACGGTACTGTTGAAGTGAAGACGGTTCATCTGATCGTTAACGGTAAAAAATCCGTG 124

QY 557 TTTCTGCTGAACGCGAACCAAGCAACATTTGACTGSGTACTGATGGCGTAGAAATCGTTC 616

Db 125 TTACCGCTGAACGCGATCCGGCTAACCTGAAATGGG-ACGAATTGGTGTGACGTAGTG 183

QY 617 TTGAAGCAACTAGTTTCTTCTGCTAAAGAAAGCTGTGAAACACACTTACATGCTAACG 676

Db 184 CTGAAGCTACCGGTATCTTCTGACTGACGAAACCGCGGTAAACACATCACCGCT---G 240

QY 677 GTGCTAAAAAAGTTGTTTATCACAGCTCCTGGTGGAAACGACGTTAAACACAGTTGTTTCA 736

Db 241 GCGGAAAAAAGTGGTTCTGACGGGTCCGTCTAAAGACAAACACCCCAATGTTTGTAAAG 300

QY 737 ACATAACCAACGACATTTCTGACGGTACTGAAACACAGTTATCTCAGGTGCTTCATGACTA 796

Db 301 GCGCTAACTTTGACAAATACGAAGGCCAGGACATCGTT---TCCAACGCTTCTCTGACCA 357

QY 797 CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTTCAGATGCAATTTGGTATCCAAAAGGTC 856

Db 358 CCAACTGCTTGGCGCGCTGGCTAAAGTTATCAACGACAACTTCGGCATCATCGAAGGTC 417

QY 857 TTATGACTACAATCCACGCTTATCTGTTGACCAATGATCTTTCAGCGGACCAACCGTG 916

Db 418 TGATGACTACTGTTTACCGGCTACTCCGCAACCCAGAAACCGTTGACGGCCCGCTCTACA 477

QY 917 GTGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976

Db 478 AA---GACTGGCGCGCGCGCGCGCGCGCTCTCAGAACATCATCCCCTCTCTACTGGCG 534

QY 977 CTGCTAAAGCTATCGGTCTTGTATCCAGAAATGATGGTAAACTTTGATGGTGTGCTGCTG 1036

Db 535 CTGCTAAAGCGGTAGGTAAAGTAACTGCTGCGGAACTGAAATGGCAAACTGACTGGTATGCG 594

QY 1037 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACCTTTGATAAAAACG 1096

Db 595 TCCGCGTCTCTACTCCGAAACGTTATCGTTGTTGACCTGACCGTTCTGCTGGAAGGCGG 654

QY 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTC-----AAACGACAG 1145

Db 655 CTACCTACGAGCAGATCAAGCTGCTGTTAAAGCTGCTGCGGAAAGGAAACATGAAAGCGT 714

QY 1146 TTTCTGTTACACTGAAGATCCAATTTCTTCTTCAGATATCGGTAGGCGTGTATACGCTTC 1205

Db 715 GCTGGTTACACCGAAGACGACGCTGTTATCTACCGATTTCAACGGTGAAGTATGCACATTC 774

QY 1206 ATTGTTTACGCAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGTTAAAGT 1265

Db 775 CGTGTTCGATGCTAAAGCAGGCTATCGCGTGAACGACAC-----TTCGTGAAACT 825

QY 1266 TGTATCATGGTATGACAAATGAAATGCTTTACAC 1298

Db 826 GGTCTCTGTTGATGATGACAAACCGGTTACTC 858

RESULT 4

CK261340

LOCUS

DEFINITION

CK261340

CK261340.1 GI:39818318

EST.

CK261340

EST707418 potato abiotic stress cDNA library Solanum tuberosum cDNA

clone POAB929 5' end, mRNA sequence.

CK261340

CK261340.1

EST.

Db 185 TTGGTGTCAATGCTGAACCTTTACACCATGATGAACCTATCATCAGCAATGCCTCTTGCA 244

Qy 794 CTACAAACCTGTTAGCTCCTATGCTGCTAAAGCTCTTTCACGATGCAATTTGGTATCCAAAAG 853

Db 245 CCACCAACCTGCTGCTCTTTCGTCAGGTTCTTGACCGAGAAATTTGGAATTTATCAAGG 304

Qy 854 GTCTTATGACTACAAATCCACGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 913

Db 305 GAACAATGACAACTACTCACTCTTACACCGGTGACCAAGGCTTCTTGATGCAAGCCAC- 363

Qy 914 GTGGTGGTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 973

Db 364 -----AGGGATCTTAGACGTGCACGAGTGCAGCATTTGAACATAGTTCCAACTCACTG 418

Qy 974 GTGCTGCTTAAGCTATCGGCTCTTGTATCCAGAAATTTGAATGTTAAACTTTGATGCTGCTG 1033

Db 419 GTGCTGCTTAAGCTGCTGCTCTTGTCTCTCCCAAGCCTCAAGGGGAAACTCAATGGGATTG 478

Qy 1034 CACAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1093

Db 479 CCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 538

Qy 1094 ACGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1153

Db 539 AGACATTTGCTGAGGAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598

Qy 1154 AACTGAAGATCCAAATGTTTCTTTCAGATAT---CGTAGGCGTGCATACGGTTTCATTGT 1210

Db 599 GCATTCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 658

Qy 1211 TTGACGCAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGTTAAAGTTGTAT 1270

Db 659 CATCAACCATGATTTCTTCACTCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 718

RESULT 7

CK276937

LOCUS

DEFINITION

EST723015 potato abiotic stress cdna library Solanum tuberosum cdna

clone POADW68 5' end, mRNA sequence.

CK276937

CK276937.1 GI:39833915

EST.

KEYWORDS

SOURCE

Solanum tuberosum (potato)

ORGANISM

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE

1 (bases 1 to 955)

Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.

Generation of ESTs from abiotic stressed potato tissue

Unpublished (2003)

Other_ESTs: EST723016

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from TIGR via potato@tigr.org

Seq primer: ATT TAG GTG ACA CTA TAG.

Location/Qualifiers

1..955

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="POADW68"

/tissue_type="abiotic stress treated leaf and root tissue"

/lab_host="DH10B-Tona"

FEATURES

source

/clone lib="potato abiotic stress cdna library"

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, 1d, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots:3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cdna library. RNA sample."

ORIGIN

Query Match 13.5%; Score 181.6; DB 14; Length 955;

Best Local Similarity 55.2%; Pred. No. 1e-35;

Matches 421; Conservative 0; Mismatches 329; Indels 12; Gaps 3;

Qy 523 GGTGATTGGAAGTTAAGGTCATTTGTTAAAGTTTCTGCTGAACGCAACAGCAAAAC 582

Db 203 GATGGAATCTCTGTCGATGGAAATATCATCAAGTCGCTCTCAACCGTGACCTGTGAAC 262

Qy 583 ATTGACTGGGCTACTGATGGCGTAGAATCGTTCTTGAAGCAACTAGTTTCTTTGCTAAA 642

Db 263 CTCCCATGGGGAGAACTTGAATTTGTTGTCATAGAAGGTACCGAGTGTGTGTAGAC 322

Qy 643 AAAGAAGCTGCTGAAAAAACAACCTTACATGCTAACGGTGCTAAAAAGTTGTTATCACAGCT 702

Db 323 AGAGAAGGTGCCGTTAAACACATCCAGGC---CGGAGCCAAAGAGTGCTCATCACCGCC 379

Qy 703 CCTGGTGGAAACGACGTTAAACACAGTTGTTTCAACACTAACACGACATTTCTTGACGGT 762

Db 380 CCCGAAAAGGTGATATCCCTACTTATGTTGTTGGTGTCAATGCTGAACCTTTACAACCAT 439

Qy 763 ACTGAAAACAGTTATCTCAGGTGCTTCTCATGTACTACAAACTGTTTAGTCTCTATGGCTAAA 822

Db 440 GATGAACCTATCATCAGCAATGCCTCTTGACCAACCAACTGCCTTGTCTCTTCTGCTCAAG 499

Qy 823 GCTCTTACGATGCAATTTGGTATCCAAAAAGTCTTTATGACTACAATCCAGCTTATATCT 882

Db 500 GTTCTTGACCAAGAAATTTGGAATTTATCAAGGGAAACAATGACAACTACTCACTCTTACACC 559

Qy 883 GGTGACCAAAATGATCCTTGACGGACCAACCGGTGGTGACCTTCTGTCGTCGTCGTCGTC 942

Db 560 GGTGACCAAAAGGCTTCTTGATGCAAGCCACAGG-----GATCTTAGACGTGCACGAGCT 613

Qy 943 GGTGCTGCAAAACATTTGTTCTTAACCTCAACTGGTGTCTTAAAGCTATCGGTCTTGTATC 1002

Db 614 GCAGCATTTGAACATAGTTCCAACTCAACTGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTC 673

Qy 1003 CCAGAAATTTGAATGTTAAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062

Db 674 CCAAGCCTCAAGGGAAACTCAATGGGATTGCCCTCCGTTTCCACCCCTCAACGCTCTCA 733

Qy 1063 GTAAGTGTGTTGTTAACTCTTGTATAAAACCTTTCTGTTGACGAAATCAACGCTGCT 1122

Db 734 GTTGTGACCTTGTGCTTCAAGTCAACCAAGAAAGCATTTGCTGAGGAAGTGAATGCTGCA 793

Qy 1123 ATGAAAGCTGCTTCAAAACGACAGTTTTCGGTTTACACTGAAGATCCAATTTGTTTCTTCA 1182

Db 794 TTCAGAAAGCTGCTGTATAAGAACTCAATGGCATTTCTAGCTGTCTGCGATGAACCACTT 853

QY 1183 AT---CGTAGGGTGTATACCGTTTCATTGTTTGACCGCAACTCAAACTAAAGTTATGGAA 1239
 Db 854 GTGTCAAGTTGACTTCCGGTGCAGTGATGTGTCATCAACCAATTGATTCTTCACTACCATG 913
 QY 1240 GTTGACGGATCACAAATTGGTTAAAGTTGTATCATCATGGTATGAC 1281
 Db 914 GTCATGGGAGATGACATGGTTAAAGGTATTGCTTGGTATGAC 955

RESULT 8
 CK272330/c
 LOCUS
 DEFINITION EST718408 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POAD453 3' end, mRNA sequence.

ACCESSION CK272330.1 GI:39829308

VERSION EST.

KEYWORDS Solanum tuberosum (potato)

SOURCE Solanum tuberosum

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 891)
 AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
 TITLE Generation of ESTs from abiotic stressed potato tissue

JOURNAL Unpublished (2003)

COMMENT Other ESTs: EST718407

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from TIGR via potato@tigr.org

Seq primer: GTA ATA CGA CTC ACT ATA GGG C.

Location/Qualifiers

FEATURES

source

1. 891
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POAD453"
 /tissue_type="abiotic stress treated leaf and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="potato abiotic stress cDNA library"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation of
 the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
 and 5d). Set 3 were grown under the standard conditions
 and then were cold stressed by placement at 4 C. Cold
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d. Set 4 were grown under the standard conditions and
 then were heat stressed by placement at 35 C. Heat
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d and 4d and heat-stressed roots were harvested at 6 hr,
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and
 equal RNA from each tissue and stress was pooled to
 construct the cDNA library. RNA sample."

ORIGIN

Query Match 13.3%; Score 179.2; DB 14; Length 891;
 Best Local Similarity 55.7%; Pred. No. 4.1e-35;
 Matches 409; Conservative 0; Mismatches 313; Indels 12; Gaps 3;

QY 617 TTGAAGCAACTAGTTTCTTTGTAAAAAAGAAGCTGCTGAAAAACACTTACATGCTAACG 676
 Db 889 TAGAAGGTACCGGAGTGTGTTGTAGACAGAGAAGGTGCGGTAAACACATCCAGGC---CG 833
 QY 677 GTGCTAAAAAAGTTGTTATCACAGCTCCTGGTGGAAACGAGCTTAAAAACAGTTGTTTCA 736
 Db 832 GAGCCAGAAGGTGCTCATCACGCCCCCGGAAAAGGTGATCCCTACTTATGTTGTG 773
 QY 737 AACTAACCAACGACATTTCTTACCGGTACTGAAACAGTTATCTCAGGTGCTTCACTACTA 796
 Db 772 GTGTCAATGCTGAACCTTTACAAACCATGATGAACCTTATCATCAGCAATGCCTCTTGCACCA 713
 QY 797 CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTTCACGATGCAATTTGGTATCCAAAAGGTC 856
 Db 712 CCAACTGCTTGTCTCCTTTCGTCGAAGTTCTTTCGACCAAGAAATTTGGAATTATCAAGGGAA 653
 QY 857 TTATGACTACAAATCCACGCTTATCTGCTGACCAAAATGATCTTACGGACCAACCGTG 916
 Db 652 CAATGACAACTACTCACTCTTACACCGGTGACCAAAAGGCTTGTGATGCAAGCCAC---- 597
 QY 917 GTGGTGACCTTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
 Db 596 --AGGATCTTAGACGTGACGAGCTGCAGCATTTGAACATAGTTTCCAACTCAACTGGTG 539
 QY 977 CTGCTAAAGCTATCGGTCTTGTGTTATCCAGAAATTTGAATGTTGTTGTTGTTGTTGTTGCT 1036
 Db 538 CTGCTAAGCTGTGGCTCTTGTCTCCCAAGCCTCAAGGGGAAACTCAATGGGATTGCCCC 479
 QY 1037 AACGTGTTCTGTTTCCAACTGGAATCAGTAACAGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1096
 Db 478 TCCGTGTTCCCAACCCCTAACAGTCTCAGTTGTGGACCTTGTCGTTCAAGTCAACCAAGA 419
 QY 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGTTTCAAAACGACAGTTTTCGGTTACA 1156
 Db 418 CATTTGCTGAGGAAGTGAATGCTGTCATTCAGAGAAAGCTGCTGATAAGGAACCTCAATGGCA 359
 QY 1157 CTGAAGATCCAAATGTTTCTTCAGATAT---CGTAGGCGTGTCTACATACGGTTCATTGTTG 1213
 Db 358 TTCTAGCTGTCTGCGATGAACCACTTGTGTGTCAGTTGACTTCCGGTGCAGTGTGTCAT 299
 QY 1214 ACGCAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGATCAT 1273
 Db 298 CAACCATTGATTTCTTCACTCACCATTGATGGGAGATGACATGGTTAAGGTTATTGCTT 239
 QY 1274 GGTATGACAATGAAATGTTTACACTGCTCAACTTGTTCGTAACCTTGAGTATTTTGCAA 1333
 Db 238 GGTATGACAATGAAATGGGTTTACTCACAGAGGGTTGTTGATCTTGTGACATTTGTTGCAA 179
 QY 1334 AAATCGCTAAATAA 1347
 Db 178 ACCAGTGGAAATGA 165

RESULT 9

CK272824

LOCUS

DEFINITION

EST718902 potato abiotic stress cDNA library Solanum tuberosum cDNA

clone POAD755 5' end, mRNA sequence.

ACCESSION CK272824

VERSION CK272824.1 GI:39829802

KEYWORDS EST.

SOURCE

ORGANISM

Solanum tuberosum (potato)

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Solanum.

1 (bases 1 to 973)

Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.

Generation of ESTs from abiotic stressed potato tissue

Unpublished (2003)

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

CK272824 973 bp mRNA linear EST 12-DEC-2003
 EST718902 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POAD755 5' end, mRNA sequence.

CK272824

CK272824.1 GI:39829802

EST.

Solanum tuberosum (potato)

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Solanum.

1 (bases 1 to 973)

Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.

Generation of ESTs from abiotic stressed potato tissue

Unpublished (2003)

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ART TAG GTG ACA CTA TAG.
Location/Qualifiers

FEATURES

Source
1. .973
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAD755"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Query Match 13.2%; Score 178.2; DB 14; Length 973;
Best Local Similarity 58.3%; Pred. No. 7.7e-35;
Matches 353; Conservative 0; Mismatches 243; Indels 9; Gaps 2;
QY 523 GGTGGATTGGAAGTTAAGTCAATTTGTTAAAGTTCTGCTGAACGGGACCAAC 582
Db 378 GATGGAATCTCTGTCGATGGAAGAAATCATCCAAAGTCGTCTCCAAACCGTGAGAAC 437
QY 583 ATTGACTGGGCTACTGATGGCGTAGAAATCGTTCTTGAAGCAACTAGTTTCTTTGCTAAA 642
Db 438 CTCCATGGGGAGAACTTGGAGTTGATTGGTCTATAGAAGGTACCGAGTGTGTGTAGAC 497
QY 643 AAAGAAGCTGCTGAAAAACACTTACATGCTAAACGGTGCTGCTGCTGCTGCTGCTGCTGCT 702
Db 498 AGAGAAGGTGCGGTAAACACATCCAGGC--CGGAGCCAAAGAGGTGCTCATCACCGCC 554
QY 703 CCTGTGGAAACGACGTTAAACAGTTGTTTCAACACTAACCAACGACATCTTGACGGT 762
Db 555 CCCGAAAGGTGATATCCCTACTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 614
QY 763 ACTGAAACAGTTATCTCAGGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822
Db 615 GATGAACCTATCATCAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 674
QY 823 GCTCTTACGATGCTTGGTATCCAAAAAGGTCTTATGACTACATCCACGCTTATCT 882
Db 675 GTTCTTGACAGAAATTTGGAATATCAAAAGGAACAATGACAACTACTCATTCTTACACC 734
QY 883 GGTGACCAATGATCCTTGACGGACCAACCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 942
Db 735 GGTGACCAAGGCTTCTTGATGCAAGCCACAGGGAT-----CTTAGACGTGCAACGACT 788
QY 943 GGTGCTGCAAAACATTTCTTAACTCAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1002
Db 789 GCAGCACTCAACATAGTTCCAACTCAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 848

QY 1003 CCAGAAATTGAATGGTAAACTTGTATGGTGTGCTGCAACAGTGTCTCTTCCAACTGGATCA 1062
Db 849 CCAAGCCTTAAGGGGAAACTACGGCAATGCTCTCCGTTGTTCCACCCCTAACGTCTCT 908
QY 1063 GTAACCTGAGTTGTTGTAACCTCTTGTATATAAAAGCTTTCTGTTGACGAATCAACGCTGCT 1122
Db 909 GTTGTGGACCTTGTGCTTTCAGTCAACCAAGAGACATTTGCTGAGGAAGTGAATGCTGCA 968
QY 1123 ATGAA 1127
Db 969 TTCAA 973

RESULT 10

AI779267
LOCUS
DEFINITION
641 bp mRNA linear EST 18-MAY-2001
clone CLES7N21, mRNA sequence.

ACCESSION

AI779267.1 GI:5277308

VERSION

AI779267

KEYWORDS

EST.

SOURCE

Lycopersicon esculentum (tomato)

ORGANISM

Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 641)

REFERENCE

1 (bases 1 to 641)
D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,
Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y.,
Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,
Giovannoni, J.J. and Martin, G.B.
Generation of ESTs from Pseudomonas susceptible tomato

TITLE

Unpublished (1999)

JOURNAL

Contact: CUGI

COMMENT

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

Location/Qualifiers

1..641

source

/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="R11-13 (Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="CLES7N21"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato susceptible, Cornell"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; CLES - Tomato Pseudomonas Susceptible EST Library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site"

ORIGIN

Query Match 13.2%; Score 177.8; DB 9; Length 641;
Best Local Similarity 58.1%; Pred. No. 8.4e-35;
Matches 355; Conservative 0; Mismatches 247; Indels 9; Gaps 2;
QY 523 GGTGGATTGGAAGTTAAGTCAATTTGTTAAAGTTTCTGCTGAACGGGACCAAC 582
Db 23 GATGGCATCTCTGTCGATGGAGAAAGTCATCCAAGTCGTCTCCAACCGTGACCCCTGTGAAC 82
QY 583 ATTGACTGGGCTACTGATGGCGTAGAAATCGTTCTTGAAGCAACTAGTTTCTTTGCTAAA 642
Db 83 CTCCATGGGGAGAACTTGGAGTTGACTTAGTCATAGAAGGTACCGGAGTTTGTGTAGAC 142
QY 643 AAAGAAGCTGCTGAAAAACACTTACATGCTAACGGTGTGCTAAAAAGTTGTTATCACAGCT 702
Db 143 AGAGAAGGTGCGGTAAACACATCCAGGC--CGGAGCCAAAGAGGTGCTCATCACCGCC 199

ORIGIN		Mixed callus was harvested at 25 and 40 days and included undifferentiated masses. Tomato Callus EST Library"									
	Query Match	13.2%;	Score 177.8;	DB 12;	Length 783;						
	Best Local Similarity	58.1%;	Pred. No. 9e-35;								
	Matches 355;	Conservative 0;	Mismatches 247;	Indels 9;	Gaps 2;						
QY	523	GGTGGATTGGAAGTTAACGGTCAATTGTTAAAGTTTCTGCTGAACGCGAACCGCAAC	582								
Db	73	GATGGCATCTCTGTCGATGGAAAAAGTCATCCAAGTCGTCTCCAACCGTGACCCCTGTGAAC	132								
QY	583	ATTGACTGGGCTACTGATGGCGGTAGAAATCGTTCITGAAGCAAACTAGTTTCTTTTGCTAAA	642								
Db	133	CTCCCATGGGAGAACCTGGAGTTGACTTAGTCATAGAAGGTACCGGAGTTTTTTGTAGAC	192								
QY	643	AAAGAAGCTGCTGAAAAACACTTACATGCTAACGGGTGCTAAAAAAGTTGTTATCACAGCT	702								
Db	193	AGAGAAGGTGCCGGTAAACACATCCAGGC--CGGAGCCAAGAAGGTGCTCATCACCGCC	249								
QY	703	CCTGGTGAAACGACGTTAAAAACAGTTGTTTTCAACACTAACCAACGACATTCCTTGACGGT	762								
Db	250	CCCGGAAAAGGTGACATCCCTACTTATGTTGTTGGTGTCATGCTGAACCTTTACAGCCAT	309								
QY	763	ACTGAAACAGTTATCTCAGGTGCTTCATGTACTACAAACTGTTTAGCTCCTATGGCTAAA	822								
Db	310	GATGAACCTATCATCAGCAATGCCTCTTGTAACCAACTGCCTTGCTCCTTTCGTCAAG	369								
QY	823	GCTCTTCACGATGCAATTTGGTATCCAAAAAGGTCTTTATGACTACAATCCACGCTTATACT	882								
Db	370	GTTCTTGACCAGAAATTTGGAATTATCAAGGGAACAATGACAACACTACTCACTCTTACACC	429								

Strain	Sequence	Position
Db	GGTGACCAAAAGGCTTCTTGATGCAAGCCACAGGGAT-----CTTAGACGTGCACGAGCT	483
Qy	GGTGCTGCAAAACATTGTTCCCTAACTCAACTGGTGGTGTGCTGCTAAAGCTATCGGTCTTTGTTATC	1002

484	DB	GCAGCACTGAACATAGTTCCTCAACCTCAACTGGTGGCTGCTAAAGGCTGTGGCTCTTGTCTC	543
1003	QY	CCAGAATTGAATGGTAACTTGATGGTGTGCACAAACGTTCTCTGTTCCAACTGGATCA	1062
544	DB	CCAAGCCTCAAGGGGAAACTCAACGGCATTGCCCTCCGTGTTCCCAACCCCTAACGTCCTCG	603

[illegible]

Qy	1123	ATCAAAGCTGC	1133
Db	664	TTCAGAGAGGC	674

RESULT 12

AW267163					
LOCUS	AW267163	530 bp	mRNA	linear	EST 30-DEC-1999
DEFINITION	YA61 gastric carcinoma cell GC7901 Homo sapiens cDNA, mRNA sequence.				

ACCESSION	AW267163
VERSION	AW267163.1
KEYWORDS	GI:6647328
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 530)
AUTHORS	Zhao,J.R., Yan,X.J., Han,F.C., Cui,D.X., Hou,Y., Yan,Q.J. and Su,C.Z.
TITLE	Gastric associated differentially expressed gene mRNA sequence
JOURNAL	Unpublished (1999)
COMMENT	Contact: Zhao JR Institute of Genetic Diagnosis, the Fourth Military Medical University


```
RESULT 14
BQ045826      771 bp      mRNA      linear      EST 10-MAR-2003
LOCUS
DEFINITION
P. infestans-challenged potato leaf, incompatible
reaction Solanum tuberosum cDNA clone BPL113A7 5' end, mRNA
sequence.
BQ045826      771 bp      mRNA      linear      EST 10-MAR-2003
BQ045826      771 bp      mRNA      linear      EST 10-MAR-2003
BQ045826.1    GI:19819812
EST.
Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 771)
Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukolanov, A.,
Rangel, P., Haberlach, G.T., Karamycheva, S.A., Tsai, J., Chiemiango, A.,
Bougri, O., Buell, C.R., Renning, C.M., Helgeson, J. and Baker, B.
Generation of ESTs from Potato Leaves Challenged with Phytophthora
infestans, incompatible interaction (2002)
Unpublished (2002)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.
Location/Qualifiers
1..771
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="BPL113A7"
/tissue_type="leaf"
/dev_stage="6 week old"
/lab_host="SOLR"
/clone_lib="P. infestans-challenged potato leaf,
incompatible reaction"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: UC Berkeley, PGEC; sequencing: The
Institute for Genomic Research. Whole plants were
challenged with 450,000 sporangia/ml P. infestans isolate
US-1 (US940501) in Biotron (Madison, Wisconsin). Leaf
tissue was collected at 1, 2, 5, 12, and 24 hours
post-challenge and frozen in liquid nitrogen immediately
upon removal. Kennebec plants showed no signs of
Katahdin plants (susceptible to P. infestans US-1) were
used as controls and showed infection. NOTE: We cannot
exclude the possibility that this sequence is actually
derived from Phytophthora rather than potato."
ORIGIN
Query Match      12.9%; Score 174; DB 12; Length 771;
Best Local Similarity 55.3%; Pred. No. 8.6e-34;
Matches 426; Conservative 0; Mismatches 330; Indels 14; Gaps 4;
QY 523 GGTGATTGGAAGTTAACGGTCAATTGTTAAAGTTTCTGCTGAACCGCAACAGCAAC 582
D 6 GATGGAATCTCTGTCGATGGAATAATCATCAAGTCGTCTCCAACCGTGACCTGTGAAC 65
QY 583 ATTGACTGGGCTACTGATGGCGGTAGAAATCGTTCTTGAAGCAACATGTTTCTTTGCTAAA 642
D 66 CTCCCATGGGAGAACTTGGAGTTGATTGTCATAGAAGGTACCGAGTGTGTTAGAC 125
QY 643 AAAGAAGCTGCTGAAACACTTACATGCTAACGGTCTGCTAAAGTTGTTATCACAGCT 702
D 126 AGAGAAGGTGCCGGTAAACACATCCAGGC---CGGAGCCAAAGAGGTGCTCATCACCGCC 182
QY 703 CCTGGTGGAAACGACGTTAAACAGTTGTTTCAACACTAACCCAGCATTTCTTGACGGT 762
```

```
Db 183 CCCGAAAAGGTGATATCCCTACTATGTTGTTGGTGTCAATGCTGAACCTTTACACCAT 242
QY 763 ACTGAAACAGTTATCTCAGGTGCTTCACTACTACAAAGTCTTCTCTATGGCTAAA 822
Db 243 GATGAACCTATCATCAGCAATGCCTCTTGACACCACCACTGCTTCTCTTCTGTCAG 302
QY 823 GCTCTTCACGATGCAATTTGGTATCCAAAAGGTCTTATAGCTACAAATCCACGCTTAT 882
Db 303 GTTCTTGACCAAGAAATTTGGAATTTATCAAGGAACAAATGACAACTACTCACTCTT 362
QY 883 GGTGACCAATGATCCTTGACGGACCAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 942
Db 363 GGTGACCAAGAGCTTCTTGATGCAAGCCACA-----GGGATCTTAGACGTGCA 416
QY 943 GGTGCTGCAAAACATTTGTTCCCTAACTCAACTGGTGGTGGTGGTGGTGGTGGTGGT 1002
Db 417 GCAGCACTCAACATAGTTCCCAACCTCAACTGGTGGTGGTGGTGGTGGTGGTGGT 476
QY 1003 CCAGAATTGAATGTAACCTTGAATGCTGCTGCACAAACGTTGTTCTGTTCCAACT 1062
Db 477 CCAAGCCTTAAAGGAACTCAACGGCATGCTCTCCGTTGTTCCACCCCTAACGTC 536
QY 1063 GTAACCTGAGTTGGTTGTAACCTCTTGATAA-AAACGTTTCTGTTGACGAAATCA 1121
Db 537 GTTGTGGACCTTGTGTTCAAGTCACCAAGAGACATTTTGTGAGGAAGTGAATGCT 596
QY 1122 TATGAAAGCTGCTTCAACACGACAGTTTGGTTTACACTGAAGATCCAAATGTTCT 1181
Db 597 ATTCAGAGAGGCTGCTGATAAGGAACCTCAATGGCATTCTATCTGTTGTGATGA 656
QY 1182 TATCGTAGGCGGTG-----CATACGGTTTCATTTTGACGCAACTCAACTAAAG 1237
Db 657 CGTGTCAAGTTGACTTCCGGTGCAGTGCATCAACCATGATGTTCTTCACTCACA 716
QY 1238 AAGTTGACGGATCACAATTGGTTAAAGTTTATCATGTTATGACAAATGAA 1287
Db 717 TGGTCATGGGAGATGACATGTTAAGTTATTGTTGTTGTTGTTGTTGTTGTTGTT 766
```

```
RESULT 15
LOCUS
DEFINITION
MCS038C08_160608 Ice plant Lambda Uni-Zap XR expression library, 5
days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase I (2
AM). Mesembryanthemum crystallinum cDNA clone MCS038C08 5, mRNA
sequence.
ACCESSION
CA835550
VERSION
CA835550.1 GI:26563315
KEYWORDS
EST.
SOURCE
Mesembryanthemum crystallinum (common iceplant)
ORGANISM
Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Aizoaceae; Mesembryanthemum.
REFERENCE
1 (bases 1 to 698)
AUTHORS
Cushman, J.C.
TITLE
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
JOURNAL
Unpublished (1997)
COMMENT
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T3 20mer
BACKWARD: T7 21mer
Plate: 038 row: C column: 08
Seq primer: T3 20mer
High quality sequence stop: 698.
```

FEATURES		Location/Qualifiers	
source		1..698	
		/organism="Mesembryanthemum crystallinum"	
		/mol_type="mRNA"	
		/db_xref="taxon:3544"	
		/clone="MCS038C08"	
		/tissue_type="leaf"	
		/dev_stage="five-week-old"	
		/clone_lib="Ice plant Lambda Uni-Zap XR expression library, 5 days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase I (2 AM)."	
		/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Library construction was performed according to Stratagene's recommended protocol for the Lambda UnizapXR vector and cDNA synthesis kit."	
ORIGIN			
Query Match		12.8%; Score 172.8; DB 14; Length 698;	
Best Local Similarity		57.0%; Pred. No. 1.7e-33;	
Matches		359; Conservative 0; Mismatches 262; Indels 9; Gaps 2;	
QY	514	GTAAAGAGGTTGGAATTCGAAGTTAACGGTCAATTTGTTAAAGTTTCTGCTGAACGGAA	573
Db	50	GTGGGGATGATGGCATCTCAGTTGATGGCAAAATCATCAAAGTTGTCTCAGACAGGAAC	109
QY	574	CCAGCAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTCTTGAAGCAACTAGTTTC	633
Db	110	CCTGTCAACCTTCCCTGGAAAGAAATGGGCATCGATTGGTGATAGAGGGACCGGAGTG	169
QY	634	TTTGCTAAAAAGAGCTGCTGTAAGAACACCTTACATGCTAACGGTGCTTAAAGTTGTT	693
Db	170	TTTGTGACAGAGACGGTGCAGGTAAAGCATTTCGAAGC---CGAGCTAAGAAGGTCCTC	226
QY	694	ATCAGAGCTCCTGGTGGAAACGACGTTAAACAGTTGTTTTCACACACTAACCCACGACATT	753
Db	227	ATAAAGTCTCCAGGAAAGGGTGATATTCCAACCTATGTGGTTGGTGTCATGAGGAGAAG	286
QY	754	CTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCACTACTACAAACTGTTTAGCTCCT	813
Db	287	TATGACCCAGCTGAGACCATCATCAGTAACGCCCTCCTGCACCACTAACTGCTTGGCACCC	346
QY	814	ATGGCTAAAGCTCTTACAGATGCAATTTGGTATCCAAAAAGGTCTTATGACTACAATCCAC	873
Db	347	TTTGTCAAGGTCCTCGACCAAGAAATTTGGCATCATCAAGGGAACCATGACCACTACCCAC	406
QY	874	GCTTATACCTGGTGACCAATGATCCTTGACGGACACACCGTGGTGGTGACCTTCGTCGT	933
Db	407	TCCTACACTGGTGACCAAGGCTTCTAGACGCAAGTCACC-----GTGACCTAAGGCGT	460
QY	934	GCTCGTCTGGTGTGCAAAACATTTCTTCACTCACTGGTGTGCTGCTAAAGCTATCGGT	993
Db	461	GCAAGAGCTGCAGCCCTCAACATTTGTCCTCCACCTCACTGCTGCTGCTGCTGCTGCTGCT	520
QY	994	CTTGTATCCCCAGAAATTGAATGTTAAACTTGTATGTTGCTGCACAAACGTTTCTGTTCCA	1053
Db	521	TTGGTTCTGCCCCAGCCTCAAGGGCAAGCTCAACGGGATCGCTCTTCGTGTCCTGACCCCC	580
QY	1054	ACTGATCAGTAACTGAGTTGGTTGTAACCTTTGATATAAAACGTTTCTGTTGACGAAATC	1113
Db	581	AATGCTCAGTGGTTGACCTTGTGTCTCCAGGTCAACCAAGAAGACCTTCGCTGAGGAGGTG	640
QY	1114	AACGCTGCTATGAAAGCTGCTTCAACAGAC	1143
Db	641	AATGCTGCCTTCAGAGAGGCTGCTGACAAAC	670

Search completed: March 31, 2004, 01:55:47
Job time : 3354 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 15:57:18 ; Search time 96.4776 Seconds
(without alignments)
1312.028 Million cell updates/sec

Title: US-10-650-369-22
Perfect score: 2278
Sequence: 1 MKKITGIILLLLAVILSAC.....EMSYTAQLVRLTLEYFAKIAK 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		Match	Length	DB	ID	Description
	Score	%					
1	2278	100.0	448	5	AAM50664	Aam50664	Streptoco
2	1656.5	72.7	336	5	AAM50665	Aam50665	Streptoco
3	1656.5	72.7	336	5	AAM50639	Aam50639	Streptoco
4	1655.5	72.7	336	5	ABP29960	Abp29960	Streptoco
5	1655.5	72.7	336	6	ABU46455	Abu46455	Protein e
6	1655.5	72.7	345	5	ABP29106	Abp29106	Streptoco
7	1652.5	72.5	336	2	AAR56486	Aar56486	Plasmin r
8	1652.5	72.5	336	4	AAV85681	Aay85681	Streptoco
9	1564.5	68.7	336	5	ABP30758	Abp30758	Streptoco
10	1559.5	68.5	336	5	AAM50667	Aam50667	Streptoco
11	1559.5	68.5	336	5	AAM50641	Aam50641	Streptoco
12	1557.5	68.4	336	5	AAM50666	Aam50666	Streptoco
13	1557.5	68.4	336	5	AAM50640	Aam50640	Streptoco
14	1535	67.4	335	6	ABU02516	Abu02516	S. pneumo
15	1535	67.4	359	4	AAU37576	Aau37576	Streptoco
16	1535	67.4	359	6	ABU46262	Abu46262	Protein e
17	1534.5	67.4	336	5	AAM50669	Aam50669	Streptoco
18	1534.5	67.4	336	5	AAM50643	Aam50643	Streptoco
19	1534	67.3	359	4	AAU38000	Aau38000	Streptoco
20	1530.5	67.2	336	5	AAM50668	Aam50668	Streptoco
21	1530.5	67.2	336	5	AAM50642	Aam50642	Streptoco
22	1521	66.8	337	6	ABU44330	Abu44330	Protein e
23	1491	65.5	333	2	AAW55089	Aaw55089	Streptoco
24	1491	65.5	333	5	ABP54583	Abp54583	S. pneumo
25	1491	65.5	333	7	ADC45135	Adc45135	S. pneumo

26	1357.5	59.6	336	5	ABB55601	Abb55601	Lactococc
27	1282	56.3	333	4	AAU35254	Aau35254	Enterococ
28	1282	56.3	333	6	ABU29370	Abu29370	Protein e
29	1275	56.0	336	5	ABB53868	Abb53868	Lactococc
30	1270	55.8	333	7	ADC95142	Adc95142	E. faeciu
31	1224.5	53.8	335	6	ABU25342	Abu25342	Protein e
32	1205.5	52.9	336	5	ABB48810	Abb48810	Listeria
33	1205.5	52.9	336	6	ABU32568	Abu32568	Protein e
34	1203	52.8	334	6	ABU23595	Abu23595	Protein e
35	1166	51.2	335	6	ABU43810	Abu43810	Protein e
36	1145	50.3	334	6	ABU37737	Abu37737	Protein e
37	1141	50.1	357	6	ABP81005	Abp81005	N. gonorr
38	1141	50.1	357	6	ABU37077	Abu37077	Protein e
39	1138.5	50.0	336	4	AAG83058	Aag83058	S. epider
40	1138.5	50.0	336	6	ABU42966	Abu42966	Protein e
41	1138.5	50.0	348	5	ABP40668	Abp40668	Staphyloc
42	1132.5	49.7	336	4	AAU36947	Aau36947	Staphyloc
43	1132.5	49.7	336	4	AAU34122	Aau34122	Staphyloc
44	1132.5	49.7	336	6	ABJ19195	Abj19195	Pathogen
45	1132.5	49.7	336	6	ABU16073	Abu16073	Protein e

ALIGNMENTS

RESULT 1

AAM50664
ID AAM50664 standard; protein; 448 AA.

XX AAM50664;

AC
XX
DT 29-AUG-2003 (revised)
DT 08-APR-2002 (first entry)
XX

DE Streptococcus GapC multiepitope fusion PolyGap4.

XX
KW PolyGap4; GapC; plasmin binding protein; epitope; infection; vaccine;
KW immunisation; mastitis; therapy.

OS Streptococcus dysgalactiae.

OS Streptococcus agalactiae.

OS Streptococcus parauberis.

OS Chimeric.

XX
PN WO200196379-A2.

XX
PD 20-DEC-2001.

XX
PF 11-JUN-2001; 2001WO-CA000836.

XX
PR 12-JUN-2000; 2000US-0211247P.

XX
PA (UWSA-) UNIV SASKATCHEWAN.

XX
PI Potter AA, Perez-Casal J, Fontaine M;

XX
DR WPI; 2002-098051/13.

XX
N-PSDB; ABA91327.

XX Novel GapC multiple epitope fusion polypeptide comprising antigenic determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S. parauberis, or S. iniae GapC protein useful for treating mastitis in vertebrates.

PS Claim 9; Fig 6A-C; 116pp; English.

XX The present sequence is that of a novel multiple epitope fusion protein, designated PolyGap4, comprising the entire amino acid sequence of the Streptococcus dysgalactiae GapC plasmin binding protein in addition to unique amino acid sequences from the Streptococcus parauberis and Streptococcus agalactiae GapC proteins. The multiple epitope protein is produced in host cells transformed with an expression vector comprising a chimeric gene (see ABA91327) encoding the protein. PolyGal4 is an example

CC of novel GapC multiple epitope fusion proteins of the invention that
CC comprise epitopes from 1 or more of *S. dysgalactiae*, *S. agalactiae*, *S.*
CC *parauberis*, *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-
CC 69). The multiple epitope fusion proteins are used in claimed vaccines
CC for treating or preventing a bacterial infection in a vertebrate,
CC especially a streptococcal infection, and particularly mastitis. They are
CC also used in claimed methods of detecting *Streptococcus* antibodies. The
CC multiple epitope protein is capable of eliciting broad immunity against a
CC variety of streptococcal infections while minimising the number of
CC antigens present in the final formulation and concomitantly reducing
CC production costs. (Updated on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 448 AA;

Query Match 100.0%; Score 2278; DB 5; Length 448;
Best Local Similarity 100.0%; Pred. No. 1.8e-180;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKITGIILLLLAVIILSACQANYGSMVVKVINGFGRIGRLAFRRIONVEGVETRIN 60
Db 1 MKKITGIILLLLAVIILSACQANYGSMVVKVINGFGRIGRLAFRRIONVEGVETRIN 60
QY 61 DLTDPNMLAHLKLYDTTQGRFDGTVEVKEGFEVNGNFIKVSARDPENIDWATDGEIV 120
Db 61 DLTDPNMLAHLKLYDTTQGRFDGTVEVKEGFEVNGNFIKVSARDPENIDWATDGEIV 120
QY 121 LEALEGTVEVKDGGFDVNGKFIKVSAREKDPQIDWATDGEIVLEIDGTVEVKEGFEVN 180
Db 121 LEALEGTVEVKDGGFDVNGKFIKVSAREKDPQIDWATDGEIVLEIDGTVEVKEGFEVN 180
QY 181 GQFVKVSAEREPANIDWATDGEIVLEATSFPAKKEAAEKHLHANGAKKVVITAPGGNDV 240
Db 181 GQFVKVSAEREPANIDWATDGEIVLEATSFPAKKEAAEKHLHANGAKKVVITAPGGNDV 240
QY 241 KTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHAYTGDMIL 300
Db 241 KTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHAYTGDMIL 300
QY 301 DGPHRGGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTGVSVELV 360
Db 301 DGPHRGGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTGVSVELV 360
QY 361 TLDKNVSVDEINAAKAAASNDSEFGYTEDPIVSSDIVGSYGLFDATQTKVMEVDGSQLV 420
Db 361 TLDKNVSVDEINAAKAAASNDSEFGYTEDPIVSSDIVGSYGLFDATQTKVMEVDGSQLV 420
QY 421 KVVSWYDNEMSYTAQLVRLTLEYFAKIAK 448
Db 421 KVVSWYDNEMSYTAQLVRLTLEYFAKIAK 448

RESULT 2
AAM50665
ID AAM50665 standard; protein; 336 AA.
XX
AC AAM50665;
XX
DT 08-APR-2002 (first entry)
XX
DE Streptococcus dysgalactiae gapC plasmin binding protein DysGapC.
XX
KW DysGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
XX immunisation; mastitis; therapy.
OS Streptococcus dysgalactiae.
XX
PN WO200196379-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000836.
XX
PR 12-JUN-2000; 2000US-0211247P.

XX (UYSA-) UNIV SASKATCHEWAN.
PA Potter AA, Perez-Casal J, Fontaine M;
XX
XX WPI; 2002-098051/13.
DR N-PSDB; ABA91328.
DR
XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of *Streptococcus dysgalactiae*, *S. agalactiae*, *S. uberis*, *S.*
PT *parauberis*, or *S. iniae* GapC protein useful for treating mastitis in
PT vertebrates.
XX
PS Claim 8; Fig 1A-B; 116pp; English.

XX The present sequence is that of the GapC plasmin binding protein,
CC DysGapC, of *Streptococcus dysgalactiae* ATCC 43078, an isolate from a case
CC of bovine mastitis. The invention relates to novel GapC multiple epitope
CC fusion proteins that comprise epitopes from 1 or more of *Streptococcus*
CC *dysgalactiae*, *Streptococcus agalactiae*, *Streptococcus parauberis*,
CC *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-69). A claimed
CC example is PolyGap4 (see AAM50664). Expression vectors and host cells for
CC production of the multiple epitope fusion proteins are provided. The
CC multiple epitope proteins are used in claimed vaccines for treating or
CC preventing a bacterial infection in a vertebrate, especially a
CC streptococcal infection, and particularly mastitis. They are also used in
CC claimed methods of detecting *Streptococcus* antibodies. The multiple
CC epitope proteins are capable of eliciting broad immunity against a
CC variety of streptococcal infections while minimising the number of
CC antigens present in the final formulation and concomitantly reducing
CC production costs
XX
SQ Sequence 336 AA;

Query Match 72.7%; Score 1656.5; DB 5; Length 336;
Best Local Similarity 79.6%; Pred. No. 4.9e-129;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETRINDLDPNMLAHLKLYDTTQGRFDGTVEV 87
Db 1 MVKVGINGFGRIGRLAFRRIONVEGVETRINDLDPNMLAHLKLYDTTQGRFDGTVEV 60
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVSAE 147
Db 61 KEGGFEVNGNFIKVSARDPE-----
QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 82 -----NIDWATDGEIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTN 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDMILDGPHRGGDLRRARAGAAINVPNSTGAA 327
Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDMILDGPHRGGDLRRARAGAAINVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTTLDKNVSVDEINAAKAAASNDSEFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTTLDKNVSVDEINAAKAAASNDSEFGYTE 275
QY 388 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 3
AAM50639

ID AAM50639 standard; protein; 336 AA.
XX AAM50639;
AC
XX
DT 04-APR-2002 (first entry)
XX Streptococcus dysgalactiae gapC plasmin binding protein.
DE
XX
KW GapC; plasmin-binding protein; DysgalGapC; infection; mastitis; vaccine;
KW diagnosis; therapy.
XX
OS Streptococcus dysgalactiae.
XX
PN WQ200196381-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000838.
XX
PR 12-JUN-2000; 2000US-0211022P.
XX
PA (UUSA-) UNIV SASKATCHEWAN.
XX
PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
XX
XX WPI; 2002-130725/17.
DR N-PSDB; ABA91248.
XX
XX Novel isolated GapC protein of Streptococcus dysgalactiae, S.agalactiae,
PT S.uberis, S.paraberis, or S.iniae, useful as vaccine component for
PT treating streptococcal infection which causes mastitis in vertebrates.
XX
PS Claim 1(a); Fig 1A-B; 107pp; English.
XX
XX The present sequence is that of the GapC plasmin binding protein
CC (DysGalGapC) of Streptococcus dysgalactiae ATCC 43078, a clinical isolate
CC from bovine mastitis. It is encoded by the GapC gene given in ABA91248.
CC
CC GapC protein, which has no signal sequence or membrane anchor domain, is
CC capable of eliciting an immune response in a vertebrate. The invention
CC provides the GapC genes and proteins of 5 Streptococcus species, as well
CC as recombinant vectors, host cells and vaccine compositions comprising
CC GapC polynucleotides or proteins. The vaccines are used to treat or
CC prevent a bacterial infection, especially a streptococcal infection, and
CC mastitis in particular (claimed). GapC proteins are also used in claimed
CC methods for detecting GapC antibodies, and to raise antibodies that are
CC used in claimed methods for detecting GapC proteins
XX
SQ Sequence 336 AA;

Query Match 72.7%; Score 1656.5; DB 5; Length 336;
Best Local Similarity 79.6%; Pred. No. 4.9e-129;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFEGTVEV 87
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFEGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVSAR 147
DB 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVVEIVLE 207
DB 82 -----NIDWATDGVVEIVLE 95
QY 208 ATSFPAKKEAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDILDTETVISGASCTTN 267
DB 96 ATGFPAKKEAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDILDTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGDLRRRAGAANIVPNSGTAA 327
DB 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGDLRRRAGAANIVPNSGTAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLTKNVSVDEINAAAMKAASNDSFGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLTKNVSVDEINAAAMKAASNDSFGYTE 275
QY 388 DPVSSDIVGVSGLFDATQTKMVEVDGSLVKVSVWYDNMSYTAQLVRLTLEYFAKIA 447
DB 276 DPVSSDIVGVSGLFDATQTKMVEVDGSLVKVSVWYDNMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
DB 336 K 336
RESULT 4
ABP29960
ID ABP29960 standard; protein; 336 AA.
XX
AC ABP29960;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 9096.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WQ200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-CH004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
DR N-PSDB; ABN70591.
XX
PT New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
PS Claim 1; Page 4031; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), AEN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 336 AA;

Db 336 K 336

RESULT 6

ABP29106

ID ABP29106 standard; protein; 345 AA.

XX

AC ABP29106;

XX

DT 02-JUL-2002 (first entry)

XX

DE Streptococcus polypeptide SEQ ID NO 7388.

XX

KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KW group A streptococcus; Streptococcus pyogenes; antibacterial;

KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX

OS Streptococcus pyogenes.

XX

PN WO200234771-A2.

XX

PD 02-MAY-2002.

XX

PF 29-OCT-2001; 2001WO-GB004789.

XX

PR 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

XX

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX

PI Telford J, Massignani V, Margarit y RosI, Grandi G, Fraser C;

PI Tettelin H;

XX

DR WPI; 2002-352536/38.

DR N-PSDB; ABN69737.

XX

PT New Streptococcus protein for the treatment or prevention of infection or

PT disease caused by Streptococcus bacteria, such as meningitis, and for

PT detecting a compound that binds to the protein.

XX

PS Claim 1; Page 3888; 4525pp; English.

XX

CC The invention relates to a protein (ABP25413-ABP30895) from group B

CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

CC the specification. The proteins have antibacterial and antiinflammatory

CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and

CC antibodies that bind (I) are used in the manufacture of medicaments for

CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound binds to

CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be

CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (I) may be used to recombinantly produce (I) and may be

CC used in gene therapy. Antibodies to (I) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins

XX

SQ Sequence 345 AA;

Query Match 72.7%; Score 1655.5; DB 5; Length 345;

Best Local Similarity 79.3%; Pred. No. 6.1e-129;

Matches 334; Conservative 1; Mismatches 1; Indels 85; Gaps 1;

Qy 28 MVKVGINGFRIGRLAFRIQNVGEVETRIINDLPNLAHLKLYDTTQGRFDGTVEV 87

Db 10 MVKVGINGFRIGRLAFRIQNVGEVETRIINDLPNLAHLKLYDTTQGRFDGTVEV 69

Qy 88 KEGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVSAR 147

Db 70 KEGFEVNGNFIKVSARDPE----- 90

Qy 148 XDPEQIDWATDGEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207

Db 91 -----NIDWATDGEIVLE 104

Qy 208 ATFFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTWVFNTNHDILDTGTETVISGASCTTN 267

Db 105 ATGFFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTWVFNTNHDILDTGTETVISGASCTTN 164

Qy 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGGHRRARAGAAINVPNSTGAA 327

Db 165 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGGHRRARAGAAINVPNSTGAA 224

Qy 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLDKNVSVDEINAAKKAASNDSFGYTE 387

Db 225 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLDKNVSVDEINAAKKAASNDSFGYTE 284

Qy 388 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447

Db 285 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 344

Qy 448 K 448

Db 345 K 345

RESULT 7

AAR56486

ID AAR56486 standard; protein; 336 AA.

XX

AC AAR56486;

XX

DT 25-MAR-2003 (revised)

DT 22-FEB-1995 (first entry)

XX

DE Plasmin receptor.

XX

KW Plasmin receptor; isolate 64/14; plasmin; alpha-2-antiplasmin;

KW plasminogen activator; bleeding; reocclusion; thrombosis;

KW pulmonary embolism; clots.

XX

OS Streptococcus pyogenes.

XX

PN US5328996-A.

XX

PD 12-JUL-1994.

XX

PF 10-AUG-1992; 92US-00928462.

XX

PR 29-MAR-1989; 89US-00330849.

PR 16-MAY-1990; 90US-00524411.

XX

PA (UYFL) UNIV FLORIDA RES FOUND INC.

XX

PI Von Mering G, Broder C, Boyle MDP, Lottenberg R;

XX

DR WPI; 1994-225327/27.

DR N-PSDB; AAQ70705.

XX

PT New DNA encoding bacterial plasmin receptor - useful as thrombolytic

PT agents, used with plasminogen activator or bound to plasmin, also useful

PT in vaccines against bacterial infection.

XX

PS Claim 1; Col 27-30; 19pp; English.

XX

CC This sequence represents the S. pyogenes plasmin receptor. The DNA

CC encoding this sequence was isolated from the S. pyogenes clinical isolate

CC 64/14. The plasmin receptor has a very high affinity for plasmin which,

CC when bound, retains its enzymatic activity but is not regulated

CC (inhibited) by alpha-2-antiplasmin. The receptor protein, when

CC administered concurrently or sequentially, prolongs the activity of

CC plasminogen activator (PA) so allows a reduction in dose, and thus lowers
CC the risk of bleeding, and may prevent reocclusion of blood vessels. The
CC protein may be coupled to a fibrin-specific monoclonal antibody to
CC provide targeting to clots. The plasmin receptor may be useful in human
CC or veterinary medicine, for treatment of thrombosis and pulmonary
CC embolism, and for solubilising clots in catheters or shunts. (Updated on
CC 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 336 AA;

Query Match 72.5%; Score 1652.5; DB 2; Length 336;
Best Local Similarity 79.1%; Pred. No. 1e-128;
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLTDPNMLAHLKYDTTQGRFDGTV 87
Db 1 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLTDPNMLAHLKYDTTQGRFDGTV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 61 KEGGFVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLE 207
Db 82 -----NIDWATDGVLE 95
QY 208 ATSFPAKKEAAEKHLHANGAKKVITAPGGNDVKTVEVNTNHDILDGTETVISGASCTTN 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVEVNTNHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAGFIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAIVPNSTGAA 327
Db 156 CLAPMAKALHDAGFIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVTLTKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVTLTKVSVWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 8
AA85681 ID AAY85681 standard; protein; 336 AA.
XX AC AAY85681;
XX DT 13-FEB-2001 (first entry)
XX DE Streptococcal plasmin receptor amino acid sequence.
XX KW Plasmin receptor; plr; immune response; vaccination; antibacterial;
XX KW Streptococcal infection; antihelminthic.
XX OS Streptococcus pyogenes.
XX PN US6136323-A.
XX PD 24-OCT-2000.
XX PF 11-JUL-1994; 94US-00273247.
XX PR 29-MAR-1989; 89US-00330849.
XX PR 16-MAY-1990; 90US-00524411.
XX PR 10-AUG-1992; 92US-00928462.
XX PA (UYFL) UNIV FLORIDA RES FOUND INC.

XX Von Mering G, Broder C, Lottenberg R, Boyle MDP;
PI WPI; 2001-006210/01.
XX N-PSDB; AAC66140.
DR Raising an immune response in a mammal, especially for vaccination
XX against group A streptococcal infections, comprises administering an
PT isolated streptococcal plasmin receptor protein.
PT
XX Claim 2; Col 27-30; 17pp; English.
PS Polynucleotide sequence AAC66140 encodes the Streptococcal plasmin
XX receptor protein AAY85681. The protein is used in a method for raising an
CC immune response in a mammal. The method comprises administering the
CC plasmin receptor protein. The method is useful as a vaccination against
CC group A Streptococcal infections and potentially against a broad range of
CC infections associated with pathogens expressing glyceraldehyde-3-
CC phosphate dehydrogenase (GAPDH) or GAPDH-like proteins at their cell
CC surfaces. The vaccine has antibacterial and antihelminthic activity
XX
SQ Sequence 336 AA;

Query Match 72.5%; Score 1652.5; DB 4; Length 336;
Best Local Similarity 79.1%; Pred. No. 1e-128;
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLTDPNMLAHLKYDTTQGRFDGTV 87
Db 1 MVVKVINGFGRIGRLAFRIQNVGEVTRINDLTDPNMLAHLKYDTTQGRFDGTV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLE 207
Db 82 -----NIDWATDGVLE 95
QY 208 ATSFPAKKEAAEKHLHANGAKKVITAPGGNDVKTVEVNTNHDILDGTETVISGASCTTN 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVEVNTNHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAGFIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAIVPNSTGAA 327
Db 156 CLAPMAKALHDAGFIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAAIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVTLTKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSVTELVTLTKVSVWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 9
ABP30758 ID ABP30758 standard; protein; 336 AA.
XX AC ABP30758;
XX DT 02-JUL-2002 (first entry)
XX DE Streptococcus polypeptide SEQ ID NO 10692.
XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;
XX KW antinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus agalactiae.
OS
XX WO200234771-A2.
FN
XX 02-MAY-2002.
PD
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
XX 27-OCT-2000; 2000GB-00026333.
PR
XX 24-NOV-2000; 2000GB-00028727.
PR
XX 07-MAR-2001; 2001GB-00005640.
PR
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI; 2002-352536/38.
DR N-PSDB; ABN71389.
XX
XX New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
XX Claim 1; Page 4179; 4525pp; English.
PS
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 336 AA;

Query Match 68.7%; Score 1564.5; DB 5; Length 336;
Best Local Similarity 74.1%; Pred. No. 2.1e-121;
Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTV 87
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRF 54
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVVEKDGDFVNGKFIKVS 147
DB 55 ----- 54
QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGV 207
DB 55 -----DGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 95
QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGGNDVKIVFNTNHDILDGTETVISCASCTTN 267
DB 96 ATGFFASKEAEQHIHENGAKKVVITAPGGNDVKIVFNTNHDILDGTETVISCASCTTN 155
QY 268 CLAPMAKALHDAFGIQGLMTTIHAYTGDQMLDGPGRGGDLRRARAGANIVPNSTGAA 327
DB 156 CLAPMAKALQDNFQVKGQGLMTTIHAYTGDQMLDGPGRGGDLRRARAGANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLDKNVSVDEINAAAMKAASNDSFGYTE 387

DB 216 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLEKDVVEEVNAAAMKAANDSYGYTE 275
QY 388 DPIVSSDIVGVSGLSFLDQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
DB 276 DPIVSSDIVGVSGLSFLDQTKVQTVDGNQLVKVSWYDNEMSYTSOLVRLTLEYFAKIA 335
QY 448 K 448
DB 336 K 336
RESULT 10
AAM50667
ID AAM50667 standard; protein; 336 AA.
XX
AC AAM50667;
XX
DT 08-APR-2002 (first entry)
XX
DE Streptococcus uberis gapC plasmin binding protein UberGapC.
XX
KW UberGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
KW immunisation; mastitis; therapy.
XX
OS Streptococcus uberis.
XX
XX WO200196379-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000836.
XX
PR 12-JUN-2000; 2000US-0211247P.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Potter AA, Perez-Casal J, Fontaine M;
XX
XX WPI; 2002-098051/13.
DR N-PSDB; ABA91330.
XX
PT Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
PT parauberis, or S. iniae GapC protein useful for treating mastitis in
PT vertebrates.
XX
PS Claim 8; Fig 3A-B; 116pp; English.
XX
CC The present sequence is that of the GapC plasmin binding protein,
CC UberGapC, of Streptococcus uberis ATCC 9927. The invention relates to
CC novel GapC multiple epitope fusion proteins that comprise epitopes from 1
CC or more of Streptococcus dysgalactiae, Streptococcus agalactiae,
CC Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae
CC (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664).
CC Expression vectors and host cells for production of the multiple epitope
CC fusion proteins are provided. The multiple epitope proteins are used in
CC claimed vaccines for treating or preventing a bacterial infection in a
CC vertebrate, especially a streptococcal infection, and particularly
CC mastitis. They are also used in claimed methods of detecting
CC Streptococcus antibodies. The multiple epitope proteins are capable of
CC eliciting broad immunity against a variety of streptococcal infections
CC while minimising the number of antigens present in the final formulation
CC and concomitantly reducing production costs
XX
SQ Sequence 336 AA;

Query Match 68.5%; Score 1559.5; DB 5; Length 336;
Best Local Similarity 73.9%; Pred. No. 5.5e-121;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLKYDTTQGRFDGTV 87

Db 1 M V V K V G I N G F G R I G R L A F R R I Q N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 60
QY 88 K E G G F E V N G N F I K V S A E R D P E N I D W A T D G V E I V L E A L E G T V E V K D G G F V N G K F I K V S A E 147
Db 61 K D G G F E V N G N F I K V S A E K D P E ----- 81
QY 148 K D P E Q I D W A T D G V E I V L E I D G T V E V K E G G F E V N G Q F V K V S A E R E P A N I D W A T D G V E I V L E 207
Db 82 -----N I D W A T D G V E I V L E 95
QY 208 A T S F F A K E A A E K H L H A N G A K K V I T A P G G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 267
Db 96 A T G F F A K A A A E K H L H A N G A K K V I T A P G G D D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 155
QY 268 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A A N I V P N S T G A A 327
Db 156 C L A P M A K A L Q D N F G V K Q G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A S N I V P N S T G A A 215
QY 328 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 387
Db 216 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V A V L E K E T S V E E I N A A M K A A A N D S Y G Y T E 275
QY 388 D P I V S S D I G M A Y G S L F D A T Q T K V M E V D G S Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 447
Db 276 D P I V S S D I G M A Y G S L F D A T Q T K V Q T V D G N Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 335
QY 448 K 448
Db 336 K 336

RESULT 11
AAM50641
ID AAM50641 standard; protein; 336 AA.
XX AAM50641;
AC AAM50641;
XX 04-APR-2002 (first entry)
DT Streptococcus uberis gapC plasmin binding protein.
DE Streptococcus uberis gapC protein; infection; mastitis; vaccine;
XX GapC; plasmin-binding protein; Streptococcus uberis.
KW diagnosis; therapy.
XX Streptococcus uberis.
OS WO200196381-A2.
XX 20-DEC-2001.
PN 11-JUN-2001; 2001WO-CA000838.
XX 12-JUN-2000; 2000US-0211022P.
PR (UYSA-) UNIV SASKATCHEWAN.
XX Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
PI WPI; 2002-130725/17.
XX N-PSDB; ABA91250.
XX Novel isolated GapC protein of Streptococcus dysgalactiae, S. agalactiae,
PT S. uberis, S. parauberis, or S. infantiae, useful as vaccine component for
PT treating streptococcal infection which causes mastitis in vertebrates.
XX Claim 1(c); Fig 3A-B; 107pp; English.

XX The present sequence is that of the GapC plasmin binding protein
CC (UberGapC) of Streptococcus uberis ATCC 9927, encoded by the GapC gene
CC given in ABA91250. GapC, which has no signal sequence or membrane anchor
CC domain, is capable of eliciting an immune response in a vertebrate. The
CC invention provides the GapC genes and proteins of 5 Streptococcus
CC species, as well as recombinant vectors, host cells and vaccine

CC compositions comprising GapC polynucleotides or proteins. The vaccines
CC are used to treat or prevent a bacterial infection, especially a
CC streptococcal infection, and mastitis in particular (claimed). GapC
CC proteins are also used in claimed methods for detecting GapC antibodies,
CC and to raise antibodies that are used in claimed methods for detecting
CC GapC proteins. S. uberis is a common pathogen associated with mastitis in
CC cattle, horse, sheep and goat
XX Sequence 336 AA;
SQ Query Match 68.5%; Score 1559.5; DB 5; Length 336;
Best Local Similarity 73.9%; Pred. No. 5.5e-121;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 M V V K V G I N G F G R I G R L A F R R I Q N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 87
Db 1 M V V K V G I N G F G R I G R L A F R R I Q N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 60
QY 88 K E G G F E V N G N F I K V S A E R D P E N I D W A T D G V E I V L E A L E G T V E V K D G G F V N G K F I K V S A E 147
Db 61 K D G G F E V N G N F I K V S A E K D P E ----- 81
QY 148 K D P E Q I D W A T D G V E I V L E I D G T V E V K E G G F E V N G Q F V K V S A E R E P A N I D W A T D G V E I V L E 207
Db 82 -----N I D W A T D G V E I V L E 95
QY 208 A T S F F A K E A A E K H L H A N G A K K V I T A P G G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 267
Db 96 A T G F F A K A A A E K H L H A N G A K K V I T A P G G D D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 155
QY 268 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A A N I V P N S T G A A 327
Db 156 C L A P M A K A L Q D N F G V K Q G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A S N I V P N S T G A A 215
QY 328 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 387
Db 216 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V A V L E K E T S V E E I N A A M K A A A N D S Y G Y T E 275
QY 388 D P I V S S D I G M A Y G S L F D A T Q T K V M E V D G S Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 447
Db 276 D P I V S S D I G M A Y G S L F D A T Q T K V Q T V D G N Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 335
QY 448 K 448
Db 336 K 336
RESULT 12
AAM50666
ID AAM50666 standard; protein; 336 AA.
XX AAM50666;
AC AAM50666;
XX 08-APR-2002 (first entry)
DT Streptococcus agalactiae gapC plasmin binding protein AgalGapC.
XX Streptococcus agalactiae.
KW AgalGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
XX immunisation; mastitis; therapy.
OS Streptococcus agalactiae.
XX WO200196379-A2.
PN 20-DEC-2001.
XX 11-JUN-2001; 2001WO-CA000836.
XX 12-JUN-2000; 2000US-0211247P.
XX (UYSA-) UNIV SASKATCHEWAN.
XX Potter AA, Perez-Casal J, Fontaine M;
PI

XX WPI; 2002-098051/13.
DR N-PSDB; ABA91329.
XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
PT parauberis, or S. iniae GapC protein useful for treating mastitis in
PT vertebrates.
XX
PS Claim 8; Fig 2A-B; 116pp; English.
XX
CC The present sequence is that of the GapC plasmin binding protein,
CC AgalGapC, of Streptococcus agalactiae ATCC 27541. The invention relates
CC to novel GapC multiple epitope fusion proteins that comprise epitopes
CC from 1 or more of Streptococcus dysgalactiae, Streptococcus agalactiae,
CC Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae
CC (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664).
CC Expression vectors and host cells for production of the multiple epitope
CC fusion proteins are provided. The multiple epitope proteins are used in
CC claimed vaccines for treating or preventing a bacterial infection in a
CC vertebrate, especially a streptococcal infection, and particularly
CC mastitis. They are also used in claimed methods of detecting
CC Streptococcus antibodies. The multiple epitope proteins are capable of
CC eliciting broad immunity against a variety of streptococcal infections
CC while minimising the number of antigens present in the final formulation
CC and concomitantly reducing production costs
XX
SQ Sequence 336 AA;
Query Match 68.4%; Score 1557.5; DB 5; Length 336;
Best Local Similarity 73.9%; Pred. No. 8.1e-121;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVVKVGINGFGRIGRLAFRRRIQNVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 87
Db 1 MVVKVGINGFGRIGRLAFRRRIQNVEGVETRLNDLTPNMLAHLKLYDTTQGRF----- 54
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEALEGTVKVGDFVNGKFIKVSAAE 147
Db 55 ----- 54
QY 148 KDPEQIDWATDGVETVLEIDGTVEVKEGGFEVNGQFVKVSAEREPAIDWATDGVETVLE 207
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPAIDWATDGVETVLE 95
QY 208 ATSPFFAKKEAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 96 ATGFFASKEKAGQHIHENGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIIHAYTGDQMDILDCPHRGDLRRARAGAAIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIIHAYTGDQMDILDCPHRGDLRRARAGAAIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTETVLDKNVSVDEINAAKKAASNDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTSVTETVLDKNVSVDEINAAKKAASNDSFGYTE 275
QY 388 DPIVSSDIVGSYGSLEFDTATQTKMEVDGSQLVKVSVWYDNMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVGSYGSLEFDTATQTKVQTVDGNQLVKVSVWYDNMSYTSQVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336
RESULT 13
ID AAM50640
XX AAM50640 standard; protein; 336 AA.
AC AAM50640;
XX
DT 04-APR-2002 (first entry)

XX Streptococcus agalactiae gapC plasmin binding protein.
DE
XX GapC; plasmin-binding protein; AgalGapC; infection; mastitis; vaccine;
KW diagnosis; therapy.
KW Streptococcus agalactiae.
OS
XX WO200196381-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000838.
XX
PR 12-JUN-2000; 2000US-0211022P.
XX (UYSA-) UNIV SASKATCHEWAN.
PA
XX Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
PI WPI; 2002-130725/17.
XX N-PSDB; ABA91249.
DR
XX Novel isolated GapC protein of Streptococcus dysgalactiae, S. agalactiae,
PT S. uberis, S. parauberis, or S. iniae, useful as vaccine component for
PT treating streptococcal infection which causes mastitis in vertebrates.
XX
PS Claim 1(b); Fig 2A-B; 107pp; English.
XX
CC The present sequence is that of the GapC plasmin binding protein
CC (AgalGapC) of Streptococcus agalactiae ATCC 27541, encoded by the GapC
CC gene given in ABA91249. GapC, which has no signal sequence or membrane
CC anchor domain, is capable of eliciting an immune response in a
CC vertebrate. The invention provides the GapC genes and proteins of 5
CC Streptococcus species, as well as recombinant vectors, host cells and
CC vaccine compositions comprising GapC polynucleotides or proteins. The
CC vaccines are used to treat or prevent a bacterial infection, especially a
CC streptococcal infection, and mastitis in particular (claimed). GapC
CC proteins are also used in claimed methods for detecting GapC antibodies,
CC and to raise antibodies that are used in claimed methods for detecting
CC GapC proteins. S. agalactiae is a common pathogen associated with
CC mastitis in cattle, horse, sheep and goat. It also causes septicemia,
CC meningitis, bacteraemia, impetigo, arthritis, urinary tract infections,
CC abscesses, spontaneous abortion, etc
XX
SQ Sequence 336 AA;
Query Match 68.4%; Score 1557.5; DB 5; Length 336;
Best Local Similarity 73.9%; Pred. No. 8.1e-121;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVVKVGINGFGRIGRLAFRRRIQNVEGVETRLNDLTPNMLAHLKLYDTTQGRFDGTVEV 87
Db 1 MVVKVGINGFGRIGRLAFRRRIQNVEGVETRLNDLTPNMLAHLKLYDTTQGRF----- 54
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEALEGTVKVGDFVNGKFIKVSAAE 147
Db 55 ----- 54
QY 148 KDPEQIDWATDGVETVLEIDGTVEVKEGGFEVNGQFVKVSAEREPAIDWATDGVETVLE 207
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPAIDWATDGVETVLE 95
QY 208 ATSPFFAKKEAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 96 ATGFFASKEKAGQHIHENGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIIHAYTGDQMDILDCPHRGDLRRARAGAAIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIIHAYTGDQMDILDCPHRGDLRRARAGAAIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTETVLDKNVSVDEINAAKKAASNDSFGYTE 387
|||

Db 216 KAIGLVIPELNGKLDGAAQRVPVPTGTSVTELVALEKDVTEEVNAAKMAAANDSYGYTE 275

QY 388 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447

Db 276 DPIVSSDIVGVSGLFDATQTKVQTVTDGNQLVKVSWYDNEMSYTSQLVRTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 14

ABU02516

ID ABU02516 standard; protein; 335 AA.

XX

AC ABU02516;

XX

DT 23-OCT-2003 (revised)

DT 11-FEB-2003 (first entry)

XX

DE S. pneumoniae type 4 strain protein from coding region #2094.

XX

XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;

KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;

KW gene therapy; vaccine.

XX

OS Streptococcus pneumoniae; type 4 strain.

XX

PN WO200277021-A2.

XX

PD 03-OCT-2002.

XX

PF 27-MAR-2002; 2002WO-IB002163.

XX

PR 27-MAR-2001; 2001GB-00007658.

XX

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX

PI Massignani V, Tettelin H, Fraser C;

XX

DR WPI; 2003-040579/03.

DR N-PSDB; ABX07806.

XX

PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,

PT useful as medicaments for treating or preventing a disease or infection

PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or

PT ear infection.

XX

PS Claim 1; SEQ ID NO 4188; 56pp; English.

XX

CC The invention relates to a protein comprising or having at least 50%

CC identity to any of the 2469 amino acid sequences, identified in the

CC specification (available on a computer readable format), or its fragment,

CC expressed from 2469 of 2489 identified DNA coding regions from the

CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as

CC ABS56454. Also included are an antibody which binds one of the proteins,

CC treating a patient by administering the protein, DNA or antibody (in a

CC composition), a kit comprising first and second primers, which are the

CC nucleic acid cited above or fragments between nucleotides 8-100 of a

CC sequence not defined in the specification, for amplifying a target

CC sequence contained within a Streptococcus nucleic acid sequence, where

CC the first primer is substantially complementary to the target sequence

CC and the second primer is substantially complementary to the complement of

CC the target sequence, and where the parts of the primers having

CC substantial complementarity define the termini of the target sequence to

CC be amplified, assay comprising contacting a test compound with the

CC protein, and determining whether the test compound binds to the protein

CC and a Streptococcus pneumoniae bacterium, where one or more genes

CC encoding the proteins has been rendered inactive. The proteins, nucleic

CC acid molecules, antibody and compositions are useful as medicaments for

CC treating or preventing a disease or infection due to streptococcus

CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis

CC media or ear infection. They are also useful in developing vaccines,

CC diagnostics and antibiotics. The methods are useful for identifying

CC immunodominant proteins. The present sequence is one of the 2469 proteins

CC expressed by the identified coding regions from the genomic sequence.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to

CC standardise OS field)

XX

SQ Sequence 335 AA;

Query Match 67.4%; Score 1535; DB 6; Length 335;

Best Local Similarity 73.6%; Pred. No. 6e-119;

Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVVKVINGFGRIGRLAFRRIQNVGEVETRINDLTDPMMLAHLKYDTTQGRFDGIVEV 87

Db 1 MVVKVINGFGRIGRLAFRRIQNVGEVETRINDLTDPMMLAHLKYDTTQGRFDGIVEV 60

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGIVEVKGDFVNGKFIKVSAAE 147

Db 61 KEGGFEVNGKFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVIEVLEIDGIVEVKEGGFEVNGQFVKVSAREPANIDWATDGVIEVLE 207

Db 82 -----QIDWATDGVIEVLE 95

QY 208 ATSFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267

Db 96 ATGFFFAKKEAAEKHL-KGGAKKVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 154

QY 268 CLAPMAKALHDADFQIKQGLMTTIHAYTGDQMI LDGPHRGDLRRARAGAAANIVPNSTGAA 327

Db 155 CLAPMAKALQDNFQVVEGLMTTIHAYTGDQMI LDGPHRGDLRRARAGAAANIVPNSTGAA 214

QY 328 KAIGLVIPELNGKLDGAAQRVPVPTGTSVTELVTLDKNVSVDEINAAKMAANDSFYGYTE 387

Db 215 KAIGLVIPELNGKLDGSAQRVPTPTGTSVTELVAVLEKNVTVEVNAAKMAANDSFYGYTE 274

QY 388 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447

Db 275 DPIVSSDIVGVSGLFDATQTKVLDVDGKQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 334

QY 448 K 448

Db 335 K 335

RESULT 15

AAU37576

ID AAU37576 standard; protein; 359 AA.

XX

AC AAU37576;

XX

DT 14-FEB-2002 (first entry)

XX

DE Streptococcus pneumoniae cellular proliferation protein #5.

XX

KW Antisense; prokaryotic cellular proliferation protein; antibiotic;

KW antibacterial; drug design.

XX

OS Streptococcus pneumoniae.

XX

PN WO200170955-A2.

XX

PD 27-SEP-2001.

XX

PF 21-MAR-2001; 2001WO-US009180.

XX

PR 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

27-NOV-2000; 2000US-0253625P.
22-DEC-2000; 2000US-0257931P.
16-FEB-2001; 2001US-0269308P.
Search completed: March 29, 2004, 16:33:20
Job time : 99.4776 secs

PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS55435.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 13169; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 359 AA;

Query Match 67.4%; Score 1535; DB 4; Length 359;
Best Local Similarity 73.6%; Pred. NO. 6.6e-119;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
QY 28 M V K V G I N G F G R I G L A F R R I Q N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 87
Db 25 M V K V G I N G F G R I G L A F R R I Q N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 84
QY 88 K E G G F E V N G N F I K V S A E R D P E N I D W A T D G V E I V L E A L E G T V E V K D G G F D W N G K F I K V S A E 147
Db 85 K E G G F E V N G N F I K V S A E R D P E N I D W A T D G V E I V L E A L E G T V E V K D G G F D W N G K F I K V S A E 105
QY 148 K D P E Q I D W A T D G V E I V L E I D G T V E V K E G G F E V N G Q F V K V S A E R E P A N I D W A T D G V E I V L E 207
Db 106 -----Q I D W A T D G V E I V L E 119
QY 208 A T S F F A K E A E K H L H A N G A K K V I T A P G G N D V K T V V E N T N H O I L D G T E T V I S G A S C T T N 267
Db 120 A T G F F A K E A E K H L - K G G A K K V I T A P G G N D V K T V V E N T N H D V L D G T E T V I S G A S C T T N 178
QY 268 C L A P M A K A L H D A F G I Q K L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A A N I V P N S T G A A 327
Db 179 C L A P M A K A L Q D N F G V E G L M T T I H A Y T G D Q M I L D G P H R G G D L R R A R A G A A N I V P N S T G A A 238
QY 328 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V V T L D K N V S V D E I N A A K A A S N D S F G Y T E 387
Db 239 K A I G L V I P E L N G K L D G S A Q R V P T P T G S V T E L V A V L E K N V T V D E V N A A K A A S N E S Y G Y T E 298
QY 388 D P I V S S D I V G S Y G S L F D A T Q T K V M E V D G S Q L V K V V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 447
Db 299 D P I V S S D I V G M S Y G S L F D A T Q T K V L D V D G K Q L V K V V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 358
QY 448 K 448
Db 359 K 359

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 16:31:39 ; Search time 31.5224 Seconds
(without alignments)
733.715 Million cell updates/sec

Title: US-10-650-369-22
Perfect score: 2278
Sequence: 1 MKKITGIILLALLAVIILSAC.....EMSYTAQLVRLTYFAKIAK 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2278	100.0	448	4	US-09-878-766A-22
2	1656.5	72.7	336	4	US-09-878-766A-12
3	1652.5	72.5	336	1	US-07-928-462-2
4	1652.5	72.5	336	3	US-08-273-247-2
5	1559.5	68.5	336	4	US-09-878-766A-16
6	1557.5	68.4	336	4	US-09-878-766A-14
7	1534.5	67.4	336	4	US-09-878-766A-20
8	1530.5	67.2	336	4	US-09-878-766A-18
9	1491	65.5	333	3	US-08-961-083-54
10	1491	65.5	333	4	US-09-536-784-54
11	1282	56.3	346	4	US-09-134-000C-4400
12	1270	55.8	333	4	US-09-107-532A-4769
13	1138.5	50.0	348	4	US-09-134-001C-5513
14	923.5	40.5	357	4	US-09-134-000C-4229
15	918.5	40.3	340	4	US-09-634-238-233
16	905.5	39.7	349	4	US-09-489-039A-13939
17	849	37.3	335	6	5290690-9
18	838	36.8	334	6	5290690-10
19	803	35.3	340	3	US-09-095-855-205
20	803	35.3	340	4	US-09-205-426-205
21	780	34.2	334	6	5290690-11
22	776.5	34.1	335	4	US-09-489-039A-7679
23	758.5	33.3	282	4	US-09-107-532A-4057
24	747.5	32.8	340	4	US-09-543-681A-8083
25	719	31.6	352	4	US-09-489-039A-9410
26	690	30.3	336	4	US-09-198-452A-662
27	683.5	30.0	338	4	US-09-091-725-10

28	678	29.8	268	2	US-08-997-080-186	Sequence 186, App
29	678	29.8	268	2	US-08-997-362-186	Sequence 186, App
30	678	29.8	268	3	US-09-095-855-186	Sequence 186, App
31	678	29.8	268	4	US-09-324-542-186	Sequence 186, App
32	678	29.8	268	4	US-09-205-426-186	Sequence 186, App
33	675	29.6	335	2	US-08-903-800A-6	Sequence 6, Appli
34	668.5	29.3	340	4	US-09-252-991A-21612	Sequence 21612, A
35	660	29.0	334	4	US-09-674-826B-6	Sequence 6, Appli
36	646	28.4	341	4	US-09-134-001C-3783	Sequence 3783, Ap
37	628	27.6	333	3	US-09-532-803-11	Sequence 11, Appli
38	617	27.1	333	3	US-09-532-803-9	Sequence 9, Appli
39	601.5	26.4	277	4	US-09-548-938A-13	Sequence 13, Appli
40	542.5	23.8	442	4	US-09-540-236-3378	Sequence 3378, Ap
41	540.5	23.7	350	4	US-09-489-039A-11854	Sequence 11854, A
42	535	23.5	340	4	US-09-328-352-6704	Sequence 6704, Ap
43	532.5	23.4	417	4	US-09-540-236-3470	Sequence 3470, Ap
44	512.5	22.5	359	4	US-09-543-681A-6082	Sequence 6082, Ap
45	504.5	22.1	490	4	US-09-328-352-4616	Sequence 4616, Ap

ALIGNMENTS

RESULT 1
US-09-878-766A-22
; Sequence 22, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
; OTHER INFORMATION: epitope fusion protein
US-09-878-766A-22

QY	1	MKKITGIILLALLAVIILSACQANYGSGMVVKVINGFGRIGRLAFRRIONVEGVETRIN	60
Db	1	MKKITGIILLALLAVIILSACQANYGSGMVVKVINGFGRIGRLAFRRIONVEGVETRIN	60
QY	61	DLTDENMLAHLKDYDTTQGRFDGTVEVKEGGFVNGNFVKVSAERDPENIDWATDGVETV	120
Db	61	DLTDENMLAHLKDYDTTQGRFDGTVEVKEGGFVNGNFVKVSAERDPENIDWATDGVETV	120
QY	121	LEALEGTVEVKDGGFDVNGKFKVSAEKDPEQIDWATDGVETVLEIDGTVEVKEGGFEVN	180
Db	121	LEALEGTVEVKDGGFDVNGKFKVSAEKDPEQIDWATDGVETVLEIDGTVEVKEGGFEVN	180
QY	181	GQFVKVSAEREPANIDWATDGVETVLEATSFPAKKEAEKHLHANGAKKVVITAPGNDV	240
Db	181	GQFVKVSAEREPANIDWATDGVETVLEATSFPAKKEAEKHLHANGAKKVVITAPGNDV	240
QY	241	KTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHAYTGDMIL	300
Db	241	KTVVFNTHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHAYTGDMIL	300
QY	301	DGPHRGDLRRRAGAANIVPNTSGAAGAIGLVIPELNGKLDGAAQRPVPTGVTETLVV	360

Db 301 DCPHRRGDLRRARAGAAIVPNSTGAAGAIGLVIPBLNGKLDGAAQRPVPTGVTSLV 360
QY 361 TLDKNVSVDEINAAKAAASNDSPGYTDPVSSDIVGVSYSGLFDATQTKMEVDGSQLV 420
Db 361 TLDKNVSVDEINAAKAAASNDSPGYTDPVSSDIVGVSYSGLFDATQTKMEVDGSQLV 420
QY 421 KVVSWYDNEMSYTAQLVRLTLEYFAKIAK 448
Db 421 KVVSWYDNEMSYTAQLVRLTLEYFAKIAK 448

RESULT 2
US-09-878-766A-12
; Sequence 12, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 12
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-09-878-766A-12

Query Match 72.7%; Score 1656.5; DB 4; Length 336;
Best Local Similarity 79.6%; Pred. No. 5e-148;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 61 KEGGFVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV 207
Db 82 -----NIDWATDGVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTETV 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTETV 155

QY 268 CLAPMAKALHDAFGIQGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAIVPN 327
Db 156 CLAPMAKALHDAFGIQGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAIVPN 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLVTLVTLVTLVTLVTLV 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLVTLVTLVTLVTLV 275

QY 388 DPTVSSDIVGVSYSGLFDATQTKMEVDGSQLVKNVSWYDNEMSYTAQLVRL 447
Db 276 DPTVSSDIVGVSYSGLFDATQTKMEVDGSQLVKNVSWYDNEMSYTAQLV 335

QY 448 K 448
Db 336 K 336

RESULT 3
US-07-928-462-2
; Sequence 2, Application US/07928462
; Patent No. 5328996

GENERAL INFORMATION:
; APPLICANT: Boyle, Michael D.P.
; APPLICANT: Lottenberg, Richard
; APPLICANT: Broder, Christopher C.
; APPLICANT: von Mering, Gregory O.
; TITLE OF INVENTION: Bacterial Plasmid Receptors as
; TITLE OF INVENTION: Fibrinolytic Agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,462
; FILING DATE: 19920810
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/524,411
; FILING DATE: 16-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/330,849
; FILING DATE: 29-MAR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-928-462-2

Query Match 72.5%; Score 1652.5; DB 1; Length 336;
Best Local Similarity 79.1%; Pred. No. 1.2e-147;
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 61 KEGGFVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV 207
Db 82 -----NIDWATDGVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTETV 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTETV 155

QY 268 CLAPMAKALHDAFGIQGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAIVPN 327
Db 156 CLAPMAKALHDAFGIQGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAIVPN 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLVTLVTLVTLVTLVTLV 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLVTLVTLVTLVTLV 275

QY 388 DPVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
Db 276 DPVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 4
US-08-273-247-2
; Sequence 2, Application US/08273247
; Patent No. 6136323
; GENERAL INFORMATION:
; APPLICANT: Boyle, Michael D.P.
; APPLICANT: Lottenberg, Richard
; APPLICANT: Broder, Christopher C.
; APPLICANT: von Mering, Gregory O.
; TITLE OF INVENTION: Bacterial Plasmin Receptors as
; TITLE OF INVENTION: Fibrinolytic Agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,247
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/524,411
; FILING DATE: 16-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/330,849
; FILING DATE: 29-MAR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-273-247-2

Query Match 72.5%; Score 1652.5; DB 3; Length 336;
Best Local Similarity 79.1%; Pred. No. 1.2e-147;
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAB 147
Db 61 KEGGFVNGNFIKVSARDPE-----81
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAREPANIDWATDGVLE 207
Db 82 -----NIDWATDGVLE 95

QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTN 267
Db 96 ATGFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVVTLDKXSVSDEINAAKAAASNDSEFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVVTLDKXSVSDEINAAKAAASNDSEFGYTE 275
QY 388 DPVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
Db 276 DPVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 5
US-09-878-766A-16
; Sequence 16, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-766A-16

Query Match 68.5%; Score 1559.5; DB 4; Length 336;
Best Local Similarity 73.9%; Pred. No. 7.1e-139;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAB 147
Db 61 KEGGFVNGNFIKVSABKDPE-----81
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAREPANIDWATDGVLE 207
Db 82 -----NIDWATDGVLE 95

Query Match 72.5%; Score 1652.5; DB 3; Length 336;
Best Local Similarity 79.1%; Pred. No. 1.2e-147;
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLDPNMLAHLKDYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAB 147
Db 61 KEGGFVNGNFIKVSARDPE-----81
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAREPANIDWATDGVLE 207
Db 82 -----NIDWATDGVLE 95

Db	276	DPIVSSDIIGMAYGSLFDATQTKVQTVTDGNQLVKVSVWYDNMSYTAQLVRTLEYFAKIA	335
Qy	448	K 448	
Db	336	K 336	
RESULT 6			
US-09-878-766A-14			
; Sequence 14, Application US/09878766A			
; Patent No. 6660270			
; GENERAL INFORMATION:			
; APPLICANT: Potter, Andrew A.			
; APPLICANT: Perez-Casal, Jose			
; APPLICANT: Fontaine, Michael			
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN			
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION			
; FILE REFERENCE: 9000-0057			
; CURRENT APPLICATION NUMBER: US/09/878,766A			
; CURRENT FILING DATE: 2001-09-10			
; NUMBER OF SEQ ID NOS: 22			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 14			
; LENGTH: 336			
; TYPE: PRT			
; ORGANISM: Streptococcus agalactiae			
US-09-878-766A-14			
Query Match 68.4%; Score 1557.5; DB 4; Length 336;			
Best Local Similarity 73.9%; Pred. No. 1.1e-138;			
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;			
Qy	28	MVVKVGINGFGRIGRLAFRRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV	87
Db	1	MVVKVGINGFGRIGRLAFRRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRF	54
Qy	88	KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVS	147
Db	55		54
Qy	148	KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE	207
Db	55		95
Qy	208	ATSPFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN	267
Db	96	ATGFFASKEKAGQHIHENGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN	155
Qy	268	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAIVPNSTGAA	327
Db	156	CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAIVPNSTGAA	215
Qy	328	KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKKAASNDSFGYTE	387
Db	216	KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKKAASNDSFGYTE	275
Qy	388	DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNMSYTAQLVRTLEYFAKIA	447
Db	276	DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNMSYTAQLVRTLEYFAKIA	335
Qy	448	K 448	
Db	336	K 336	
RESULT 7			
US-09-878-766A-20			
; Sequence 20, Application US/09878766A			
; Patent No. 6660270			
; GENERAL INFORMATION:			
; APPLICANT: Potter, Andrew A.			
; APPLICANT: Perez-Casal, Jose			
; APPLICANT: Fontaine, Michael			
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN			
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION			
; FILE REFERENCE: 9000-0057			
; CURRENT APPLICATION NUMBER: US/09/878,766A			
; CURRENT FILING DATE: 2001-09-10			
; NUMBER OF SEQ ID NOS: 22			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 14			
; LENGTH: 336			
; TYPE: PRT			
; ORGANISM: Streptococcus agalactiae			
US-09-878-766A-14			
Query Match 68.4%; Score 1557.5; DB 4; Length 336;			
Best Local Similarity 73.9%; Pred. No. 1.1e-138;			
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;			
Qy	28	MVVKVGINGFGRIGRLAFRRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV	87
Db	1	MVVKVGINGFGRIGRLAFRRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRF	54
Qy	88	KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVS	147
Db	55		54
Qy	148	KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE	207
Db	55		95
Qy	208	ATSPFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN	267
Db	96	ATGFFASKEKAGQHIHENGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN	155
Qy	268	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAIVPNSTGAA	327
Db	156	CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAIVPNSTGAA	215
Qy	328	KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKKAASNDSFGYTE	387
Db	216	KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKKAASNDSFGYTE	275
Qy	388	DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNMSYTAQLVRTLEYFAKIA	447
Db	276	DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSVWYDNMSYTAQLVRTLEYFAKIA	335
Qy	448	K 448	
Db	336	K 336	
RESULT 8			
US-09-878-766A-18			
; Sequence 18, Application US/09878766A			
; Patent No. 6660270			
; GENERAL INFORMATION:			
; APPLICANT: Potter, Andrew A.			
; APPLICANT: Perez-Casal, Jose			
; APPLICANT: Fontaine, Michael			
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN			
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION			
; FILE REFERENCE: 9000-0057			
; CURRENT APPLICATION NUMBER: US/09/878,766A			
; CURRENT FILING DATE: 2001-09-10			
; NUMBER OF SEQ ID NOS: 22			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 18			
; LENGTH: 336			
; TYPE: PRT			
; ORGANISM: Streptococcus parauberis			
US-09-878-766A-18			
Query Match 67.2%; Score 1530.5; DB 4; Length 336;			
Best Local Similarity 73.4%; Pred. No. 3.9e-136;			

Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2;
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFD----- 55
QY 88 KEGGFVNGNFIVKSAERDPENIDWATDGVIEALEGTVKVGKGGFDVNGKFIKVS 147
Db 56 -----GTVEVKDGGFDVNGKFIKVS 77
QY 148 KDPEQIDWATDGVIEALEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIE 207
Db 78 KDPEQ-----IDWATDGVIE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVVFNTHDILDTGTVISGASCTTN 267
Db 96 ATGFFAKKAAAEKHLHENGAKKVVITAPGGDDVKTIVVFNTHDILDTGTVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPVHRRRARRAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVKQGLMTTHAYTGDQMLDGPVHRRRARRAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTVTELVTLDKNVSVDEINAAKAAASNDSPGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTVTELVAVLNKETSVEEINVMKAAANDSPGYTE 275
QY 388 DPVSSDIVGVSFSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPVSSDIVGVSFSLFDATQTKVQTVDGNQLVKVSVWYDNEMSYTAQLDRTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 9
US-08-961-083-54
; Sequence 54, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-54
Query Match 65.5%; Score 1491; DB 3; Length 333;
Best Local Similarity 72.4%; Pred. No. 2.1e-132;
Matches 302; Conservative 12; Mismatches 17; Indels 86; Gaps 2;
QY 29 VVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 88
Db 1 VVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 89 EGGFVNGNFIVKSAERDPENIDWATDGVIEALEGTVKVGKGGFDVNGKFIKVS 148
Db 61 EGGFVNGKFIKVS 80
QY 149 DPEQIDWATDGVIEALEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIE 208
Db 81 -----QIDWATDGVIE 95
QY 209 TSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVVFNTHDILDTGTVISGASCTTN 268
Db 96 TGFFAKKEAAEKHL-KGGAKKVVITAPGGNDVKTIVVFNTHDVLDTGTVISGASCTTN 154
QY 269 LAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPVHRRRARRAGAAANIVPNSTGAAK 328
Db 155 LAPMAKALQDNFVVEGLMTTHAYTGDQMLDGPVHRRRARRAGAAANIVPNSTGAAK 214
QY 329 AIGLVIPELNGKLDGAAQRPVPTGTVTELVTLDKNVSVDEINAAKAAASNDSPGYTED 388
Db 215 AIGLVIPELNGKLDGSAQRPVPTGTVTELVAVLNKVTVDVNAAMKAAASNESYGYTED 274
QY 389 PIVSSDIVGVSFSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAK 445
Db 275 PIVSSDIVGVSFSLFDATQTKVLDVDGKQLVKVSVWYDNEMSYTAQLVRLTGLRK 331

RESULT 10
US-09-536-784-54
; Sequence 54, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-536-784-54

Query Match 65.5%; Score 1491; DB 4; Length 333;
Best Local Similarity 72.4%; Pred. No. 2.1e-132;
Matches 302; Conservative 12; Mismatches 17; Indels 86; Gaps 2;
29 VVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEVK 88
1 VVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEVK 60
89 EGGFEVNGNFIKVSARDPENIDWATDGVIEVLEAEGTVEVKDGGFDVNGKFIKVSAAEK 148
61 EGGFEVNGKFIKVSARDPE-----80
149 DPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLEA 208
81 -----QIDWATDGVIEVLEA 95
209 TSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTTFVNTNHDILDTGTTETVISGASCTTNC 268
96 TGFFAKKEAAEKHL-KGAKKVVITAPGGNDVKTTFVNTNHDVLDGTETVISGASCTTNC 154
269 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGA 328
155 LAPMAKALQDNFVVEGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGA 214
329 AIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLTKNVSVDENIAMAAMKAAASND 388
215 AIGLVIPELNGKLDGSAQRPVPTGTSVTELVTLTKNVSVDENIAMAAMKAAASND 274
389 PIVSSDIVGVSGLSFLDTATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAK 445
275 PIVSSDIVGMSYGLSFLDTATQTKVLDVDGKQLVKVSVWYDNEMSYTAQLVRLTLEYFAK 331

RESULT 11
US-09-134-000C-4400
Sequence 4400, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: Patent in version 3.1
SEQ ID NO 4400
LENGTH: 346
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-4400

Query Match 56.3%; Score 1282; DB 4; Length 346;
Best Local Similarity 62.5%; Pred. No. 1.2e-112;
Matches 262; Conservative 25; Mismatches 46; Indels 86; Gaps 3;
28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87
14 MTVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 73
88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEAEGTVEVKDGGFDVNGKFIKVSAAE 147

Db 74 HEGSFNVNGKEIKVLNRNPEELPW-----98
QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 99 -----GEL-----GVDIVLE 108
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTTFVNTNHDILDTGTTETVISGASCTTN 267
Db 109 CTGFFTSKEAAEKHLTA-GAKRVVISAPGGNDVPTIYNTNHTLTGEEVVISGASCTTN 167
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAA 327
Db 168 CLAPMAKALHDNFGVVEGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAA 227
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLTKNVSVDENIAMAAMKAAASND 387
Db 228 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLTKNVSVDENIAMAAMKAAASND 287
QY 388 DPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLEA 446
Db 288 DEIVSSDIVGMSYGLSFLDTATQTKVLDVDGKQLVKVSVWYDNEMSYTAQLVRLTLEYFAK 346

RESULT 12
US-09-107-532A-4769
Sequence 4769, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
City: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4769:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...333
SEQUENCE DESCRIPTION: SEQ ID NO: 4769:
US-09-107-532A-4769

QY 445 KI 446
Db 356 NL 357

RESULT 15
US-09-634-238-233
; Sequence 233, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christenson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-233

Query Match 40.3%; Score 918.5; DB 4; Length 340;
Best Local Similarity 46.0%; Pred. No. 2.3e-78;
Matches 196; Conservative 49; Mismatches 88; Indels 93; Gaps 7;

QY 28 MVVKVINGFGRIQRLAPRRIONV-----EGVEVTRINDLTDPNMLAHLKYDTTQGRFDG 83
Db 1 MTVKIGINGFGRIQRLAPRRIONV-----EGVEVTRINDLTDPNMLAHLKYDTTQGRFDG 60

QY 84 TVEVKEGGFEVNGNFIKVSARDPENIDWATDGVLEIVLEALEGIVEVKDGGDFDVGKFIK 143
Db 61 EVSATDNGIVVDGKEYRVYAEPPAQNI PW-----VKV----- 92

QY 144 VSAEKDPEQIDWATDGVLEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPPANIDWATDVE 203
Db 93 -----DGVD 96

QY 204 IVLEATSFPAKKEAEKHLHANGAKKVITAPGNDVKTTFVNTNHDILDTGTETVISGAS 263
Db 97 YVLECTGFYTSAEKSOAHLDA-GAKRVLISAPAGK-IKTIVNVNDDTLNADDDKIVSAGS 154

QY 264 CTTNCLAPMAKALHDAFGIQKGLMTTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNS 323
Db 155 CTTNCLAPMAYFLNQEFGEVGTMTTTHAYTSTQMLLDGPRVGGNLRRAARSAAANTIPHS 214

QY 324 TGAAKAIGLVIPELNGKLDGAAQRPVPTGVSVELVVTLL-DKNVSVDEINAAMK--AASN 380
Db 215 TGAAKAIGLVIPELNGKLDGAAQRPVPTGVSVELVVTLL-DKNVSVDEINAAMK--AASN 274

QY 381 DSFGYTEDPIVSSDIVGSYGSFLDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTL 440
Db 275 PSFGWNEDEIVSSDIVGSYGSFLDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTL 334

QY 441 EYFAKI 446
Db 335 LKFATL 340

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 16:35:22 ; Search time 73.5522 Seconds
(without alignments)
1593.758 Million cell updates/sec

Title: US-10-650-369-22
Perfect score: 2278
Sequence: 1 MKKITGIILLALLAVIILSAC.....EMSYAQLVRLTLEYFAKIAK 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2278	100.0	448	9	US-09-878-766A-22
2	1656.5	72.7	336	9	US-09-878-766A-12
3	1656.5	72.7	336	10	US-09-878-781-4
4	1656.5	72.7	336	14	US-10-134-297-4
5	1655.5	72.7	336	12	US-10-282-122A-74379
6	1650	72.4	335	10	US-09-878-781-14
7	1562.5	68.6	336	10	US-09-878-781-16
8	1559.5	68.5	336	9	US-09-878-766A-16
9	1559.5	68.5	336	10	US-09-878-781-8
10	1559.5	68.5	336	14	US-10-134-297-8
11	1557.5	68.4	336	9	US-09-878-766A-14
12	1557.5	68.4	336	10	US-09-878-781-6
13	1557.5	68.4	336	14	US-10-134-297-6
14	1535	67.4	359	9	US-09-815-242-13169
15	1535	67.4	359	12	US-10-282-122A-74186

16	1534.5	67.4	336	9	US-09-878-766A-20	Sequence 20, Appl
17	1534.5	67.4	336	10	US-09-878-781-12	Sequence 12, Appl
18	1534.5	67.4	336	14	US-10-134-297-12	Sequence 12, Appl
19	1534	67.3	359	9	US-09-815-242-13593	Sequence 13593, A
20	1530.5	67.2	336	9	US-09-878-766A-18	Sequence 18, Appl
21	1530.5	67.2	336	10	US-09-878-781-10	Sequence 10, Appl
22	1530.5	67.2	336	14	US-10-134-297-10	Sequence 10, Appl
23	1521	66.8	337	12	US-10-282-122A-72254	Sequence 72254, A
24	1491	65.5	333	9	US-09-765-272-54	Sequence 54, Appl
25	1357.5	59.6	336	15	US-10-369-493-18524	Sequence 18524, A
26	1282	56.3	333	9	US-09-815-242-10847	Sequence 10847, A
27	1282	56.3	333	12	US-10-282-122A-57294	Sequence 57294, A
28	1280	56.2	337	15	US-10-369-493-18342	Sequence 18342, A
29	1224.5	53.8	335	12	US-10-282-122A-53266	Sequence 53266, A
30	1205.5	52.9	336	12	US-10-282-122A-60492	Sequence 60492, A
31	1203	52.8	334	12	US-10-282-122A-51519	Sequence 51519, A
32	1166	51.2	335	12	US-10-282-122A-71734	Sequence 71734, A
33	1145	50.3	334	12	US-10-282-122A-65661	Sequence 65661, A
34	1141	50.1	357	12	US-10-282-122A-65001	Sequence 65001, A
35	1138.5	50.0	336	12	US-10-282-122A-70890	Sequence 70890, A
36	1132.5	49.7	336	9	US-09-815-242-5618	Sequence 5618, Ap
37	1132.5	49.7	336	9	US-09-815-242-12540	Sequence 12540, A
38	1132.5	49.7	336	12	US-10-282-122A-43997	Sequence 43997, A
39	941.5	41.3	333	12	US-10-282-122A-56828	Sequence 56828, A
40	918.5	40.3	340	15	US-10-264-213-141	Sequence 141, App
41	916.5	40.2	340	9	US-09-971-536-65	Sequence 65, Appl
42	916.5	40.2	340	15	US-10-264-213-252	Sequence 252, App
43	908.5	39.9	332	12	US-10-282-122A-60108	Sequence 60108, A
44	900.5	39.5	336	12	US-10-282-122A-57807	Sequence 57807, A
45	870.5	38.2	337	12	US-10-282-122A-64191	Sequence 64191, A

ALIGNMENTS

RESULT 1
US-09-878-766A-22
; Sequence 22, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
; OTHER INFORMATION: epitope fusion protein
US-09-878-766A-22

Query Match	100.0%	Score 2278;	DB 9;	Length 448;
Best Local Similarity	100.0%;	Pred. No. 3.4e-199;		
Matches 448;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MKKITGIILLALLAVIILSACQANYGSGMVVKVINGFGRIGRLAFRRIQNVGVEVTRIN	60	
Db	1	MKKITGIILLALLAVIILSACQANYGSGMVVKVINGFGRIGRLAFRRIQNVGVEVTRIN	60	
Qy	61	DLTDPNMLAHLKLYDTTQGRFDGTVEVKEGFEVNGNFVKVSAERDPENIDWATDGVIEIV	120	
Db	61	DLTDPNMLAHLKLYDTTQGRFDGTVEVKEGFEVNGNFVKVSAERDPENIDWATDGVIEIV	120	
Qy	121	LEALEGTVEVKDGGFDVNGKFKVSAEKDPEQIDWATDGVIEIVLEIDGTVEVKEGGFEVN	180	

Db	121	LEALEGTVVEVKDGGFDVNGKFKIVSAEKDPEQIDWATGDGVEIVLEIDGTVEVKEGGFEVN	180
QY	181	GQFVKVSAEREPANIDWATGDGVEIVLEATSTFFAKKEAAEKHLHANGAKKVITAPGGNDV	240
Db	181	GQFVKVSAEREPANIDWATGDGVEIVLEATSTFFAKKEAAEKHLHANGAKKVITAPGGNDV	240
QY	241	KTVVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDADFQIKGLMTTIHAYTGDQMIL	300
Db	241	KTVVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDADFQIKGLMTTIHAYTGDQMIL	300
QY	301	DGPHRGGBLRRARAGAANI VPNSTGAAKAIGLVIPELNGKLDGAAQRVPVPTGVSVELVV	360
Db	301	DGPHRGGBLRRARAGAANI VPNSTGAAKAIGLVIPELNGKLDGAAQRVPVPTGVSVELVV	360
QY	361	TLDKNVSVDEINAAAMKAASNDSTFGYTEDPIVSSDIVGSVCSLFDATQTKVMEVDGSQLV	420
Db	361	TLDKNVSVDEINAAAMKAASNDSTFGYTEDPIVSSDIVGSVCSLFDATQTKVMEVDGSQLV	420
QY	421	KVSWYDNEMSYTAQLVRTLEYFAKIAK	448
Db	421	KVSWYDNEMSYTAQLVRTLEYFAKIAK	448

RESULT 2

US-09-878-766A-12

US 02 070 100R 12
: Sequence 12: Application US/09878766A

; Sequence ID, APPICACCTG 9
; Patent No. US20020044928A1

; FALCIC NO: CB20V2001
: GENERAL INFORMATION:

APPLICANT: Potter, Andrew A.

APPLICANT: PULLER, Andrew H.
APPLICANT: PEREZ-CASAL, Jose

APPLICANT: Fontaine, Michael

INVENTOR: MICHAEL J. FOLICARE, AFFIDAVIT: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN

DATE	TITLE OF INVENTION:	TITLE OF INVENTION:
1940	IMMUNIZATION OF DAIRY CATTLE WITH	AGAINST STREPTOCOCCUS INFECTION

FILE OF INVENTION: AGAT
FILE REFERENCE: 9000-0057

FILE REFERENCE: 9000-0037
CURRENT APPLICATION NUMBER: IIS/09/878-766A

; CURRENT APPLICATION NUMBER: 03/0
 : CURRENT FILING DATE: 2001-09-10

; CURRENT FILING DATE: 20
 ; NUMBER OF SEC ID NOS: 22

```

; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver 2.0

```

; SOFTWARE: F
: GEO ID NO 12

SEQ ID NO 12
LENGTH: 336

```

; LENGTH: 3
TYPE: DDE

```

TYPE: PRI
CPCNTNM: STANTON COUNTY DISTRICT

; ORGANISM: SET

Query Match	72.7%	Score 1656.5;	DB 9;	Length 336;
Best Local Similarity	79.6%	Pred. No. 1.2e-142;		
Matches 335: Conservative	0;	Mismatches 1;	Indels 85;	Gaps 1;

Qy	28	MVVKVGINGFGRIGRLAFRRIQNV	EGVEVTRINDLTDPNMLAHLKYD	TTQGRF	DGIVEV	87		
Db	1	MVVKVGINGFGRIGRLAFRRIQNV	EGVEVTRINDLTDPNMLAHLKYD	TTQGRF	DGIVEV	60		
Qy	98	KEGGFEVNGNFIKVS	AERDPENIDWATD	GVIEVLEALEG	TVKDG	GFVNGKF	IKVSAE	147
Db	61	KEGGFEVNGNFIKVS	AERDPE	-----	-----	-----	-----	81
Qy	148	KDPEQIDWATD	GVIEVLEIDG	TVKEGGFEVNGQFV	KVSAEREPANID	WATD	GVIEVLE	207
Db	82	-----	-----	-----	-----	NIDWATD	GVIEVLE	95
Qy	208	ATSPFAKKEAAEKHLHANGAKKV	VTAPGGNDVKT	VVENTNHD	LDGTET	VISG	ASCTTN	267
Db	96	ATGFFAKKEAAEKHLHANGAKKV	VTAPGGNDVKT	VVENTNHD	LDGTET	VISG	ASCTTN	155
Qy	268	CLAPNAKALHDAFGIO	KGLMTTIHAYTG	DQMILDGPHRG	DDLRRAPAGA	ANIVPN	STGAA	327
Db	156	CLAPNAKALHDAFGIO	KGLMTTIHAYTG	DQMILDGPHRG	DDLRRAPAGA	ANIVPN	STGAA	215
Qy	328	KAIGLVIPELNGKLDGAA	QRVPVPTGSVTE	LVTLTKNV	SVDEINAMKAA	SND	SFGYTE	387
Db	216	KAIGLVIPELNGKLDGAA	QRVPVPTGSVTE	LVTLTKNV	SVDEINAMKAA	SND	SFGYTE	275
Qy	388	DPIVSSDPIGVSVY	XSGLFDATQTKVME	VDGSQLVKVSV	WYDNEMSYTAQL	VRTL	EYFAKIA	447

RESULT 4

US-10-134-297-4

US-10-133-22 : Sequence 4. Application US/10134297

Publication No. US20030165524A1

GENERAL INFORMATION:

APPLICANT: Bolton, Alexandra J.

APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
TITLE OF INVENTION: STREPTOCOCCUS INFECTION
FILE REFERENCE: 9000-0055.20
CURRENT APPLICATION NUMBER: US/10/134,297
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus dysgalactiae
US-10-134-297-4

Query Match 72.7%; Score 1656.5; DB 14; Length 336;
Best Local Similarity 79.6%; Pred. No. 1.2e-142;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGFEVNGNFIKVSARDPENIDWATDVEIVLEALEGTVKGGFDVNGKFIKVS 147
Db 61 KEGFEVNGNFIKVSARDPE----- 81
QY 148 KDPEQIDWATDVEIVLEIDGTVEVKEGGFEVNGQFVKVSABREPANIDWATDVEIVLE 207
Db 82 -----NIDWATDVEIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSWYDNEMSYTAQLVRTLEYFAKIA 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSWYDNEMSYTAQLVRTLEYFAKIA 275
QY 388 DPIVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
Db 276 DPIVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 5
US-10-282-122A-74379
Sequence 74379, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 74379
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus pyogenes
US-10-282-122A-74379

Query Match 72.7%; Score 1655.5; DB 12; Length 336;
Best Local Similarity 79.3%; Pred. No. 1.5e-142;
Matches 334; Conservative 1; Mismatches 1; Indels 85; Gaps 1;
QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGFEVNGNFIKVSARDPENIDWATDVEIVLEALEGTVKGGFDVNGKFIKVS 147
Db 61 KEGFEVNGNFIKVSARDPE----- 81
QY 148 KDPEQIDWATDVEIVLEIDGTVEVKEGGFEVNGQFVKVSABREPANIDWATDVEIVLE 207
Db 82 -----NIDWATDVEIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSWYDNEMSYTAQLVRTLEYFAKIA 447
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSWYDNEMSYTAQLVRTLEYFAKIA 275
QY 388 DPIVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
Db 276 DPIVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 6
US-09-878-781-14
Sequence 14, Application US/09878781
Publication No. US20030082781A1
GENERAL INFORMATION:
APPLICANT: Bolton, Alexandra J.

```

; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SpyGapC
; OTHER INFORMATION: protein
US-09-878-781-14

Query Match      72.4%; Score 1650; DB 10; Length 335;
Best Local Similarity 79.1%; Pred. No. 4.9e-142;
Matches 333; Conservative 2; Mismatches 0; Indels 86; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDITQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDITQGRFDGTVEV 60
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 61 KEGGFEVNGNFIKVSARDPE----- 81
QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 82 -----IDWATDGEIVLE 94
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 95 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 154
QY 268 CLAPMAKALHDAPFGIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
Db 155 CLAPMAKALHDAPFGIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 214
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLVTDKNVSVDEINAAKAAASNDSEFGYTE 387
Db 215 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLVTDKNVSVDEINAAKAAASNDSEFGYTE 274
QY 388 DPIVSSDIVGSYGLFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 275 DPIVSSDIVGSYGLFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 334
QY 448 K 448
Db 335 K 335

RESULT 7
US-09-878-781-16
; Sequence 16, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16

Query Match      72.4%; Score 1650; DB 10; Length 335;
Best Local Similarity 79.1%; Pred. No. 4.9e-142;
Matches 333; Conservative 2; Mismatches 0; Indels 86; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDITQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDITQGRFDGTVEV 60
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 61 KEGGFEVNGNFIKVSARDPE----- 81
QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 82 -----IDWATDGEIVLE 94
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 95 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 154
QY 268 CLAPMAKALHDAPFGIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
Db 155 CLAPMAKALHDAPFGIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 214
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLVTDKNVSVDEINAAKAAASNDSEFGYTE 387
Db 215 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLVTDKNVSVDEINAAKAAASNDSEFGYTE 274
QY 388 DPIVSSDIVGSYGLFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 275 DPIVSSDIVGSYGLFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 334
QY 448 K 448
Db 335 K 335

RESULT 8
US-09-878-766A-16
; Sequence 16, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-766A-16

Query Match      68.5%; Score 1559.5; DB 9; Length 336;
Best Local Similarity 73.9%; Pred. No. 9e-134;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDITQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDITQGRFDGTVEV 60
```

```

; LENGTH: 336
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SeqGapC
; OTHER INFORMATION: protein
US-09-878-781-16

Query Match      68.6%; Score 1562.5; DB 10; Length 336;
Best Local Similarity 76.2%; Pred. No. 4.8e-134;
Matches 321; Conservative 2; Mismatches 13; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDITQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDITQGRFDGTVEV 60
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 61 KEGGFEVNGNFIKVSARDPE----- 81
QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 82 -----NIDWATDGEIVLE 95
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 96 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAPFGIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALHDAPFGIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAR 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLVTDKNVSVDEINAAKAAASNDSEFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLVTDKNVSVDEINAAKAAASNDSEFGYTE 275
QY 388 DPIVSSDIVGSYGLFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVGSYGLFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 8
US-09-878-766A-16
; Sequence 16, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-766A-16

Query Match      68.5%; Score 1559.5; DB 9; Length 336;
Best Local Similarity 73.9%; Pred. No. 9e-134;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDITQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTDPNMLAHLKYDITQGRFDGTVEV 60
```


RESULT 11

US-09-878-766A-14
; Sequence 14, Application US/09878766A
; Patent No. US2002004928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-766A-14

Query Match 68.4%; Score 1557.5; DB 9; Length 336;
Best Local Similarity 73.9%; Pred. No. 1.4e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVVKVGINFGFRIGRLAFRRRIQNVGVEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVGINFGFRIGRLAFRRRIQNVGVEVTRINDLDPNMLAHLKYDTTQGRF----- 54
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 55 ----- 54
QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 95
QY 208 ATSPFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTGTTETVISGASCTTN 267
Db 96 ATGFFASKEKAGQHIHENGAKKVVITAPGGNDVKTVVFNTHDILDTGTTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDMILDPHRRGGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDMILDPHRRGGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPPTGSVTELVTLTKNVSVDEINAAKKAASNDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPPTGSVTELVATLEKDVTVVEVNAAMKAAANDSYGYTE 275
QY 388 DPVSSDIVGSYGSLEFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPVSSDIVGSYGSLEFDTATQTKVTDGNQLVKVSWYDNEMSYTSQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 12

US-09-878-781-6
; Sequence 6, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781

US-09-878-781-6
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-781-6
Query Match 68.4%; Score 1557.5; DB 10; Length 336;
Best Local Similarity 73.9%; Pred. No. 1.4e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVVKVGINFGFRIGRLAFRRRIQNVGVEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVGINFGFRIGRLAFRRRIQNVGVEVTRINDLDPNMLAHLKYDTTQGRF----- 54
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 55 ----- 54
QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 95
QY 208 ATSPFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTGTTETVISGASCTTN 267
Db 96 ATGFFASKEKAGQHIHENGAKKVVITAPGGNDVKTVVFNTHDILDTGTTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDMILDPHRRGGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDMILDPHRRGGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPPTGSVTELVTLTKNVSVDEINAAKKAASNDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPPTGSVTELVATLEKDVTVVEVNAAMKAAANDSYGYTE 275
QY 388 DPVSSDIVGSYGSLEFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPVSSDIVGSYGSLEFDTATQTKVTDGNQLVKVSWYDNEMSYTSQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 13

US-10-134-297-6
; Sequence 6, Application US/10134297
; Publication No. US20030165524A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055.20
; CURRENT APPLICATION NUMBER: US/10/134,297
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-10-134-297-6
Query Match 68.4%; Score 1557.5; DB 14; Length 336;
Best Local Similarity 73.9%; Pred. No. 1.4e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVVKVGINFGFRIGRLAFRRRIQNVGVEVTRINDLDPNMLAHLKYDTTQGRFDGTVEV 87

Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRF----- 54
Qy 88 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVGVKQGGFDVNGKFIKVSAB 147
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 54
Qy 148 KDPQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 95
Qy 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVPNTNHDILDTGTETVISGASCTTN 267
Db 96 ATCFASKEKAGQIHENGAKKVVITAPGGNDVKTVPNTNHDILDTGTETVISGASCTTN 155
Qy 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDMILDPHRRGDLRRARAGAAANVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDMILDPHRRGDLRRARAGAAANVPNSTGAA 215
Qy 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKAAANDSPGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKAAANDSPGYTE 275
Qy 388 DPVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 335
Qy 448 K 448
Db 336 K 336

RESULT 14
US-09-815-242-13169
; Sequence 13169, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13169
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13169
Query Match 67.4%; Score 1535; DB 9; Length 359;

Best Local Similarity 73.6%; Pred. No. 1.7e-131;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
Qy 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRFDGTVEV 87
Db 25 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTDPVMLAHLKYDTTQGRFDGTVEV 84
Qy 88 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVGVKQGGFDVNGKFIKVSAB 147
Db 85 KEGGFEVNGKFIKVSARDPE----- 105
Qy 148 KDPQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 106 -----QIDWATDGEIVLE 119
Qy 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVPNTNHDILDTGTETVISGASCTTN 267
Db 120 ATGFFFAKKEAAEKHL-KGGAKKVVITAPGGNDVKTVPNTNHDVLDGTETVISGASCTTN 178
Qy 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDMILDPHRRGDLRRARAGAAANVPNSTGAA 327
Db 179 CLAPMAKALQDNFVGVEGLMTTHAYTGDMILDPHRRGDLRRARAGAAANVPNSTGAA 238
Qy 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKAAANDSPGYTE 387
Db 239 KAIGLVIPELNGKLDGSAQRPVPTGVSVELVAVLEKNVTVDENVAAKAAANDSPGYTE 298
Qy 388 DPVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 299 DPVSSDIVGSYSLFDATQTKVLDVVDGKQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 358
Qy 448 K 448
Db 359 K 359

RESULT 15
US-10-282-122A-74186
; Sequence 74186, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74186
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-74186

Query Match 67.4%; Score 1535; DB 12; Length 359;
Best Local Similarity 73.6%; Pred. No. 1.7e-131;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETVRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87
Db |||||||
QY 25 MVKVGINGFGRIGRLAFRRIONVEGVETVRINDLTDPVMLAHLKYDTTQGRFDGTVEV 84
Db |||||||
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db |||||||
QY 85 KEGGFEVNGKFIKVSARDPE----- 105
Db |||||||
QY 148 KDPEQIDWATDGVETVLEIDGTVEVKEGGFEVNGQFVKVSAEREPEANIDWATDGVETVLE 207
Db |||||||
QY 106 -----QIDWATDGVETVLE 119
Db |||||||
QY 208 ATSPFAKKEAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db |||||||
QY 120 ATGFFAKKEAEKHL-KGAKKVITAPGNDVKTVVFNTHDVLDTETVISGASCTTN 178
Db |||||||
QY 268 CLAPMAKALHDGFIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 327
Db |||||||
QY 179 CLAPMAKALQDNFGVVEGLMTTIHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 238
Db |||||||
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVVTLDKNVSVDEINAAMKAASNDSFGYTE 387
Db |||||||
QY 239 KAIGLVIPELNGKLDGSAQRPVPTGTSVTELVAVLEKNVTVDVNAAMKAASNESYGYTE 298
Db |||||||
QY 388 DPIVSSDIVGVSYSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
Db |||||||
QY 299 DPIVSSDIVGMSYSGLFDATQTKVLDVDGKQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 358
Db |||||||
QY 448 K 448
Db 359 K 359

Search completed: March 29, 2004, 16:45:22
Job time : 100.552 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 16:26:39 ; Search time 26.7463 Seconds
(without alignments)
1611.208 Million cell updates/sec

Title: US-10-650-369-22
Perfect score: 2278
Sequence: 1 MKKITGIILLALLAVILSAC.....EMSYTAQLVRLTYFAKIAK 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1652.5	72.5	336	2 A42963	glyceraldehyde-3-p
2	1558.5	68.4	336	2 S71350	glyceraldehyde-3-p
3	1535	67.4	335	2 F95235	glyceraldehyde 3-p
4	1535	67.4	359	2 G98099	glyceraldehyde-3-p
5	1357.5	59.6	336	2 F86905	hypothetical prote
6	1280	56.2	337	2 G86694	hypothetical prote
7	1234	54.2	334	2 S34254	glyceraldehyde-3-p
8	1205.5	52.9	336	2 AC1382	glyceraldehyde 3-p
9	1204.5	52.9	336	2 AD1751	glyceraldehyde 3-p
10	1203	52.8	334	2 C96987	glyceraldehyde 3-p
11	1145	50.3	334	2 B82019	glyceraldehyde-3-p
12	1132.5	49.7	336	2 E89850	glyceraldehyde-3-p
13	1132	49.7	334	2 E81001	glyceraldehyde-3-p
14	955.5	41.9	338	2 T09633	glyceraldehyde-3-p
15	941.5	41.3	333	2 F90881	glyceraldehyde 3-p
16	941.5	41.3	333	2 C85737	glyceraldehyde-3-p
17	870.5	38.2	337	2 S73737	glyceraldehyde-3-p
18	861.5	37.8	349	2 F90517	glyceraldehyde 3-p
19	854.5	37.5	337	2 C64233	glyceraldehyde-3-p
20	851	37.4	336	2 A43260	glyceraldehyde-3-p
21	849	37.3	335	1 DEBSG	glyceraldehyde-3-p
22	845	37.1	335	2 S12696	glyceraldehyde-3-p
23	836	36.7	335	2 H84094	glyceraldehyde-3-p
24	830	36.4	335	1 DEBSGF	glyceraldehyde-3-p
25	820.5	36.0	335	2 A70107	probable glyceral
26	819.5	36.0	333	1 DEHGGT	glyceraldehyde-3-p
27	793.5	34.8	342	2 F70391	glyceraldehyde-3-p
28	790	34.7	336	2 T36020	glyceraldehyde-3-p
29	786.5	34.5	339	2 G70915	glyceraldehyde-3-p

30	783.5	34.4	339	2 S72763	glyceraldehyde-3-p
31	781.5	34.3	334	2 AI0262	glyceraldehyde-3-p
32	780.5	34.3	331	2 F82131	glyceraldehyde 3-p
33	769.5	33.8	330	2 E75408	glyceraldehyde 3-p
34	768.5	33.7	331	2 AG0711	glyceraldehyde 3-p
35	764.5	33.6	331	1 DEECG3	glyceraldehyde-3-p
36	764.5	33.6	331	2 H90939	glyceraldehyde-3-p
37	764.5	33.6	331	2 D85788	glyceraldehyde-3-p
38	754	33.1	339	2 G64041	glyceraldehyde-3-p
39	748.5	32.9	331	1 DEUTC	glyceraldehyde-3-p
40	739	32.4	332	2 JC4373	glyceraldehyde-3-p
41	734	32.2	336	2 T40235	glyceraldehyde-3-p
42	732.5	32.2	337	2 AF2438	glyceraldehyde-3-p
43	730.5	32.1	337	2 S54141	glyceraldehyde-3-p
44	729	32.0	337	2 S40610	glyceraldehyde-3-p
45	727.5	31.9	344	2 E84043	glyceraldehyde-3-p

ALIGNMENTS

RESULT 1

A42963
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus N; Alternate names: plasmin receptor
C; Species: Streptococcus sp.
C; Date: 10-Mar-1994 #sequence revision 10-Mar-1994 #text_change 03-Jun-2002
C; Accession: A42963; B42963; JH0750
R; Lottenberg, R.; Broder, C.C.; Boyle, M.D.; Kain, S.J.; Schroeder, B.L.; Curtiss III, R
J. Bacteriol. 174, 5204-5210, 1992
A; Title: Cloning, sequence analysis, and expression in Escherichia coli of a streptococcal Reference number: A42963; MUID:92355491; PMID:1322883
A; Accession: A42963
A; Molecule type: DNA
A; Residues: 1-336 <LOT>
A; Experimental source: Group A, strain 64/14
A; Note: sequence extracted from NCBI backbone (NCBIP:110308)
A; Accession: B42963
A; Molecule type: protein
A; Residues: 2-74; 161-164, 'X', 166-174; 187-211, 'X', 213-217 <LO2>
R; Pancholi, V.; Fischetti, V.A.
J. Exp. Med. 176, 415-426, 1992
A; Title: A major surface protein on group A streptococci is a glyceraldehyde-3-phosphate A; Reference number: JH0750; MUID:92364544; PMID:1500854
A; Accession: JH0750
A; Molecule type: protein
A; Residues: 2-30, 'A', 32-40 <PAN>
C; Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C; Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F; 152/Active site: Cys #status predicted

Query Match 72.5%; Score 1652.5; DB 2; Length 336;
Best Local Similarity 79.1%; Pred. No. 9.1e-101;
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;

QY	28	MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTDNMLAHLKYDTTQGRFDGTV	87
Db	1	MVVKVINGFGRIGRLAFRRIONVEGVEVTRINDLTDNMLAHLKYDTTQGRFDGTV	60
QY	88	KEGGEVNGNFVKVSAERDENIDWATDGVIEVLEALEGTVKGGFDVNGKFKVSAE	147
Db	61	KEGGEVNGNFVKVSAERDE	81
QY	148	KDPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEV	207
Db	82	-----NIDWATDGVIEVLE	95
QY	208	ATSPFAKKEAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISCASCTTN	267
Db	96	ATGFFAKKEAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISCASCTTN	155
QY	268	CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGANIVENSTGAA	327

Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDMILDPHRRGDLRRARAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDENAMKAASNDSEFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDENAMKAASNDSEFGYTE 275

QY 388 DPIVSSDIVGSYGSFLDQATQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447

Db 276 DPIVSSDIVGSYGSFLDQATQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 2

S71350

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus

C;Species: Streptococcus "equisimilis"

C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 03-Jun-2002

C;Accession: S71350

R;Gase, K.; Gase, A.; Schirmer, H.; Malke, H.

Eur. J. Biochem. 239, 42-51, 1996

A;Title: Cloning, sequencing and functional overexpression of the Streptococcus equisimilis protein. Purification and biochemical characterization of the protein.

A;Reference number: S71350; MUID:96305364; PMID:8706717

A;Accession: S71350

A;Molecule type: DNA

A;Residues: 1-336 <GAS>

A;Cross-references: EMBL:X97788; NID:g1478268; PIDN:CAA66377.1; PID:g1478269

A;Experimental source: strain H46A

C;Genetics:

A;Gene: gapC

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C;Keywords: gluconeogenesis; glycolysis; NAD; oxidoreductase

F;4-34/Region: beta-alpha-beta NAD nucleotide-binding fold

F;152/Active site: Cys #status predicted

Query Match 68.4%; Score 1558.5; DB 2; Length 336;

Best Local Similarity 76.0%; Pred. No. 1.3e-94;

Matches 320; Conservative 2; Mismatches 14; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKDYDTQGRFDGTEV 87

Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKDYDTQGRFDGTEV 60

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVS 147

Db 61 KEGGFEVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVLEALEGTVEVKDGGFEVNGQFVKVSAEREPANIDWATDGVLE 207

Db 82 -----NIDWATDGVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTGTEVVISGASCTTN 267

Db 96 ATGFFAKKEAAEKPLHANGAKKVVITAPGGNDVKQLFSTLTTSILDGTEVVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDMILDPHRRGDLRRARAGAAANIVPNSTGAA 327

Db 156 CLAPMAKALHDAFGIQKGLMTTIHAYTGDMIVDGHRRGDLRRARAGAAANIVPNSTGAR 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDENAMKAASNDSEFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDENAMKAASNDSEFGYTE 275

QY 388 DPIVSSDIVGSYGSFLDQATQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447

Db 276 DPIVSSDIVGSYGSFLDQATQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 3

F95235

glyceraldehyde 3-phosphate dehydrogenase [imported] - Streptococcus pneumoniae (strain 7

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001

C;Accession: F95235

R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heic

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,

nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: F95235

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-335 <KUR>

A;Cross-references: GB:AE005672; PIDN:AAK76079.1; PID:g14973522; GSPDB:GN00164; TIGR:SP4

A;Experimental source: strain TIGR4

C;Genetics:

A;Gene: SP2012

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 67.4%; Score 1535; DB 2; Length 335;

Best Local Similarity 73.6%; Pred. No. 4.4e-93;

Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKDYDTQGRFDGTEV 87

Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKDYDTQGRFDGTEV 60

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVS 147

Db 61 KEGGFEVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVLEALEGTVEVKDGGFEVNGQFVKVSAEREPANIDWATDGVLE 207

Db 82 -----QIDWATDGVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTGTEVVISGASCTTN 267

Db 96 ATGFFAKKEAAEKHL-KGGAKKVVITAPGGNDVKTVVFNTHDVLDTGTEVVISGASCTTN 154

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDMILDPHRRGDLRRARAGAAANIVPNSTGAA 327

Db 155 CLAPMAKALQDNFVGVGLMTTIHAYTGDMILDPHRRGDLRRARAGAAANIVPNSTGAA 214

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDENAMKAASNDSEFGYTE 387

Db 215 KAIGLVIPELNGKLDGSAQRPVPTGVTSLVTLVAVLEKNVTVDENAMKAASNESYGYTE 274

QY 388 DPIVSSDIVGSYGSFLDQATQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447

Db 275 DPIVSSDIVGSYGSFLDQATQTKVLDVKGQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 334

QY 448 K 448

Db 335 K 335

RESULT 4

G98099

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) [imported] - St

C;Species: Streptococcus pneumoniae

C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 03-Jun-2002

C;Accession: G98099

R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: G98099
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAL00628.1; PID:g15459513; GSPDB:GN00174
C;Genetics:
A;Gene: gapA
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: oxidoreductase

Query Match 67.4%; Score 1535; DB 2; Length 359;
Best Local Similarity 73.6%; Pred. No. 4.8e-93;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETVRINDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 25 MVVKVINGFGRIGRLAFRRIONVEGVETVRINDLTPVMLAHLKYDTTQGRFDGTVEV 84

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGVEVKDGGFDVNGKFIKVS 147
Db 85 KEGGFEVNGKFIKVSARDPE----- 105

QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLE 207
Db 106 -----QIDWATDGVLE 119

QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 267
Db 120 ATGFFAKKEAEKHL-KGGAKKVVITAPGGNDVKTVVNTNHDVLDGTETVISGASCTTN 178

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANVPNSTGAA 327
Db 179 CLAPMAKALQDNFVVEGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANVPNSTGAA 238

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTSLVTLVTLKNSVDEINAAKAAASNSFGYTE 387
Db 239 KAIGLVIPELNGKLDGSAQRPVPTGSGVTSLVTLVAVLEKNVTVDEVNAAKAAASNSFGYTE 298

QY 388 DPVSSDIVGVSYSGLFDATQTKMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 299 DPVSSDIVGMSYSGLFDATQTKVLDVKGQKLVKVSWSYDNEMSYTAQLVRLTLEYFAKIA 358

QY 448 K 448
Db 359 K 359

RESULT 5
F86905
hypothetical protein gapB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: F86905
R;Bolotin, A.; Wincker, P.; Mauder, S.; Jaillon, O.; Malarne, K.; Weissensbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: F86905
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-336 <STO>
A;Cross-references: GB:AE005176; PID:g12725315; PIDN:AAK06344.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: gapB
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 59.6%; Score 1357.5; DB 2; Length 336;
Best Local Similarity 65.1%; Pred. No. 1.8e-81;
Matches 274; Conservative 21; Mismatches 41; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETVRINDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETVRINDLTPMLAHLKYDTTQGRF----- 54

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGVEVKDGGFDVNGKFIKVS 147
Db 55 ----- 54

QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLE 207
Db 55 -----DGKVEVKDGGFEVNGKFKVTAESNPANINWAEVGAIEVLE 95

QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 267
Db 96 ATGFFATKEAEKHLHANGAKKVVITAPGGSDVKTIVFNTNHEVLDGTETVISAGSCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANVPNSTGAA 327
Db 156 CLAPMADTLNKQFGIKVGTMTTVHGYTGDQMTLDGPGRGDLRRARAGAAANVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTSLVTLVTLKNSVDEINAAKAAASNSFGYTE 387
Db 216 KAIGLVIPELNGKLDGHAQRPVPTGSLTSLVTLVTLKNSVDEINAAKAAASNSFGYTE 275

QY 388 DPVSSDIVGVSYSGLFDATQTKMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPVSSDIVGINSLSLFDATQTEVTSADGAQLVTVSWYDNEMSYTSNLVRLTLEYFAKIA 335

QY 448 K 448
Db 336 K 336

RESULT 6
G86694
hypothetical protein gapA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: G86694
R;Bolotin, A.; Wincker, P.; Mauder, S.; Jaillon, O.; Malarne, K.; Weissensbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: G86694
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-337 <STO>
A;Cross-references: GB:AE005176; PID:g12723446; PIDN:AAK04657.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: gapA
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 56.2%; Score 1280; DB 2; Length 337;
Best Local Similarity 61.4%; Pred. No. 2e-76;
Matches 259; Conservative 31; Mismatches 46; Indels 86; Gaps 2;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETVRINDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETVRINDLTPMLAHLKYDTTQGRF----- 60

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGVEVKDGGFDVNGKFIKVS 147
Db 61 KEDGFDVNGKFKVTAERNPE----- 81

QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLE 207
Db 82 -----DIQWADSGVEVLE 95

QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 267
Db 96 ATGFFATKEAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISAGSCTTN 155

A;Cross-references: GB:AL592022; PIDN:CAC97780.1; PID:g16415075; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: gap
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 52.9%; Score 1204.5; DB 2; Length 336;
Best Local Similarity 58.5%; Pred. No. 1.8e-71;
Matches 247; Conservative 33; Mismatches 55; Indels 87; Gaps 4;
QY 28 MVKVGINGFGRIGRLAFRRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MTVKVGINGFGRIGRLAFRRRIQNVGEIVVAINDLTDKMLAHLKYDTTQGRFDGEVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 61 HDGFFNVNGKEVKVLANRNPEELPWG-----DL----- 88
QY 148 KQPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207
Db 89 -----GVDIVLE 95
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVPNTNHDILDGTETVISGASCTTN 267
Db 96 CTGFFTAQDKAELHIKA-GAKKVVISAPATGDMKTIYVNVNHTLDGTETVISGASCTTN 154
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
Db 155 CLAPMAKVLKDKFGVWEGMLTTIHAFTNDQNTLDAPHKGDFFRARAANAENIIPNTTGA 214
QY 328 KAIGLVPELNGKLDGAAQRPVPTGSLTELVTLDKNSVDENAMKAASN-DSFGYT 386
Db 215 KAIGEVLPKGLDGAARQRPVPTGSLTELVTLDKKTVDENVNAAMEASDPETFGYT 274
QY 387 EDPIVSSDIVGSYGSFLDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKI 446
Db 275 SDQVSSDIKGMTFGSLFDETQTKVLTVDGQQLVKTVAWYDNEMSYTAQLVRLTLEYFAKI 334
QY 447 AK 448
Db 335 AK 336

RESULT 10

C96987
glyceraldehyde 3-phosphate dehydrogenase, gene gapC [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: C96987
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: C96987
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-334 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK78686.1; PID:g15023589; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0709
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 52.8%; Score 1203; DB 2; Length 334;
Best Local Similarity 57.4%; Pred. No. 2.2e-71;
Matches 241; Conservative 41; Mismatches 52; Indels 86; Gaps 3;

QY 29 VVKVGINGFGRIGRLAFRRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRFDGTVEVK 88
Db 1 MAKTAINGFGRIGRLALRILEVPGLEVVAINDLTDKMLAHLKYDSSQGRNGEIEVK 60
QY 89 EGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVS 148

Db 61 EGAFVNGKEVKVFAEADPEKLPWG-----DL----- 87
QY 149 DPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 208
Db 88 -----GIDVVLE 95
QY 209 TSFFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVPNTNHDILDGTETVISGASCTTN 268
Db 96 TGFFTKKEAAEAHRA-GAKKVVISAPAGNDLKTIVFNVNEDLDGTETVISGASCTTN 154
QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 328
Db 155 LAPMAKVLNDKFGIEKGFMTTIHAFTNDQNTLDGPHRKGDLRRARAAAVSLIPNSTGAA 214
QY 329 AIGLVPELNGKLDGAAQRPVPTGSLTELVTLDKNSVDENAMKAASNDSFGYT 388
Db 215 AISQVIPDLAGKLDGNAQRPVPTGSLTELVSLLKKVTVVEINAMKKAADSFYGT 274
QY 389 PIVSSDIVGSYGSFLDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKI 448
Db 275 PIVSADVVGINYSGLFDTLTKIVDVNGSQLVKTAAWYDNEMSYTSQVRLTLEYFAKI 334

RESULT 11

B82019
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) C NMA0246 [impc
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C;Accession: B82019
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: B82019
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-334 <PAR>
A;Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83554.1; PID:g737900
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA0246
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: oxidoreductase

Query Match 50.3%; Score 1145; DB 2; Length 334;
Best Local Similarity 56.4%; Pred. No. 1.4e-67;
Matches 235; Conservative 35; Mismatches 61; Indels 86; Gaps 4;

QY 28 MVKVGINGFGRIGRLAFRRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MSIKVAINGFGRIGRLALRQIEKAHGEVAVVNDLTPAEMLLHLPKYDSTQGRF----- 54
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 55 -----QGTABLKDDAIWNGKEIKVFAN 77

QY 148 KQPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207
Db 78 PNPEELPW-----GEL-----GVDVVLE 95

QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVPNTNHDILDGTETVISGASCTTN 267
Db 96 CTGFFTNKTKEAHIRA-GARKVVISAPGGNDVKTVPVGVNQDILDGSETVISAASCTTN 154

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
Db 155 CLAPMAAVLQKEFGVVEGLMTTIHAYTGDQNTLDAPHKGDLLRRARAAALNIVPNSTGAA 214

QY 328 KAIGLVPELNGKLDGAAQRPVPTGSLTELVTLDKNSVDENAMKAASNDSFGYTE 387
Db 215 KAIGLVPELNGKLDGSAQRPVATGSLTELVSLEPVTKEINAMKKAASSESYCYNE 274

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 15:59:33 ; Search time 19.1045 Seconds
(without alignments)
1221.045 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278

Sequence: 1 MKKITGIILLLLAVILSAC.....EMSYTAQLVRLTLEYFAKIAK 448

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1650.5	72.5	335	1 G3P_STRPY	P50467 streptococc
2	1646.5	72.3	335	1 G3P_STRP3	Q8k8m9 streptococc
3	1553.5	68.2	335	1 G3P_STRQ	Q59906 streptococc
4	1280	56.2	337	1 G3P_LACLA	P52987 lactococcus
5	1234	54.2	334	1 G3P_CLOPA	Q59309 clostridium
6	1203	52.8	334	1 G3P_CLOAB	O52631 clostridium
7	1138.5	50.0	336	1 G3P1_STAP	Q8cpv5 staphylococ
8	1132.5	49.7	336	1 G3P1_STAM	Q925c5 staphylococ
9	955.5	41.9	338	1 G3P_LACDE	O32755 lactobacill
10	941.5	41.3	333	1 G3P3_ECOS7	P58072 escherichia
11	940.5	41.3	333	1 G3P3_ECOLI	P33898 escherichia
12	870.5	38.2	337	1 G3P_MYCPN	P75358 mycoplasma
13	854.5	37.5	337	1 G3P_MYCGE	P47543 mycoplasma
14	850	37.3	334	1 G3P_CORGL	Q01651 corynebacte
15	844	37.1	334	1 G3P1_BACSU	P09124 bacillus su
16	840	36.9	334	1 G3P_BACME	P23722 bacillus me
17	825	36.2	334	1 G3P_BACST	P00362 bacillus st
18	820.5	36.0	335	1 G3P_BORBU	P46795 borrelia bu
19	818.5	35.9	332	1 G3P_THEMA	P17721 thermotoga
20	793.5	34.8	342	1 G3P_AQUAE	O67161 aquifex aeo
21	790	34.7	336	1 G3P_STRCO	Q92518 streptomyce
22	786.5	34.5	339	1 G3P_MYCTU	O06822 mycobacteri
23	783.5	34.4	339	1 G3P_MYCLE	P46713 mycobacteri
24	777.5	34.1	339	1 G3P_MYCAV	P94915 mycobacteri
25	763.5	33.5	330	1 G3P1_SALTY	P24165 salmonella
26	759.5	33.3	330	1 G3P1_ECOLI	P06977 escherichia
27	754	33.1	339	1 G3P_HAEIN	P44304 haemophilus
28	743.5	32.6	330	1 G3P3_TRYBB	P10097 trypanosoma
29	740	32.5	332	1 G3P_RALSO	P52694 ralstonia s
30	739	32.4	332	1 G3P_STRAU	Q59800 streptomyce
31	737	32.4	337	1 G3P1_RHIRA	Q9c136 rhizomucor
32	737	32.4	337	1 G3P_MONAN	P53430 monascus an
33	734	32.2	336	1 G3P1_SCHPO	P78958 schizosacch

ALIGNMENTS

RESULT 1

G3P_STRPY

ID	G3P_STRPY	STANDARD;	PRT;	335 AA.
AC	P50467;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)			
DE	(Plasminogen-binding protein) (plasmin receptor).			
GN	GAP OR PLR OR GAPA OR SPY0274 OR SPYM18_0261.			
OS	Streptococcus pyogenes, and			
OS	Streptococcus pyogenes (serotype M18).			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1314, 186103;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-73; 160-173 AND 186-216.			
RC	STRAIN=64/14;			
RX	MEDLINE=92355491; PubMed=1322883;			
RA	Lottenberg R., Broder C.C., Boyle M.D., Kain S.J., Schroeder B.L.,			
RA	Curtiss R. III;			
RT	"Cloning, sequence analysis, and expression in Escherichia coli of a			
RT	streptococcal plasmin receptor.";			
RL	J. Bacteriol. 174:5204-5210(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SF370 / ATCC 700294 / Serotype M1;			
RX	MEDLINE=21192684; PubMed=11296296;			
RA	Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,			
RA	Primeaux C., Sezate S., Suvoirov A.N., Kenton S., Lai H.S., Lin S.P.,			
RA	Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,			
RA	Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;			
RT	"Complete genome sequence of an M1 strain of Streptococcus			
RT	pyogenes.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MGAS8232 / Serotype M18;			
RX	MEDLINE=21927593; PubMed=11917108;			
RA	Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,			
RA	Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,			
RA	Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,			
RA	Kapur V., Daly J.A., Veasy L.G., Musser J.M.;			
RT	"Genome sequence and comparative microarray analysis of serotype M18			
RT	group A Streptococcus strains associated with acute rheumatic fever			
RT	outbreaks.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).			
RN	[4]			
RP	SEQUENCE OF 1-11; 20-30; 103-128; 162-171 AND 199-215.			
RC	STRAIN=JRS4 / Serotype M6;			
RA	Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,			
RA	vanBogelen R.A.;			
RT	"Two-dimensional gel electrophoresis map of Streptococcus pyogenes			
RT	proteins.";			
RL	Submitted (MAY-2000) to Swiss-Prot.			
CC	-!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.			

34	732.5	32.2	337	1	G3P2_ANASP	P58554 anabaena sp
35	729	32.0	337	1	G3P_CLAPU	Q00584 claviceps p
36	725	31.8	331	1	G3P_THAQ	P00361 thermus aqu
37	724.5	31.8	336	1	G3P2_SYNY3	P80505 synechocyst
38	723	31.7	338	1	G3P_NEUCR	P54118 neurospora
39	722	31.7	337	1	G3P_COCHE	P29497 cochlicobolu
40	721	31.7	337	1	G3P_CURLU	P28844 curvularia
41	719.5	31.6	333	1	G3P_STRAE	P54226 streptomyc
42	717.5	31.5	330	1	G3P2_LEIME	Q01558 leishmania
43	717	31.5	335	1	G3P2_SCHPO	O43026 schizosacch
44	715	31.4	336	1	G3P_ASPNG	Q12552 aspergillus
45	713	31.3	336	1	G3P_EMENT	P20445 emericella

CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M95569; AAA26953.1; -.
CC EMBL; AE006494; AAK33348.1; -.
CC EMBL; AE009973; AAL97041.1; -.
CC HSSP; P00362; 1GD1.
CC InterPro; IPR000173; GAP_dhdrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; C; 1.
CC PRINTS; PR00078; G3PDHDRGNASE.
CC TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
FT INIT MET 0 0
FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
FT CONFLICT 261 261 A -> S (IN REF. 1).
FT SEQUENCE 335 AA; 35811 MW; F06006EE253C8A3F CRC64;
Query Match 72.5%; Score 1650.5; DB 1; Length 335;
Best Local Similarity 79.3%; Pred. No. 1.2e-93;
Matches 333; Conservative 1; Mismatches 1; Indels 85; Gaps 1;
QY 29 VVKVGINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRFDGTVEVK 88
Db 1 VVKVGINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRFDGTVEVK 60
QY 89 EGGFEVNGNFIKVSARDPENIDWATDGVETVLEALEGTVEVKDGFVDVNGKFIKVSAREK 148
Db 61 EGGFEVNGNFIKVSARDPE----- 80
QY 149 DPEQIDWATDGVETVLEIDGTVEVKEGGFEVNGQFVKVSAREPANIDWATDGVETVLEA 208
Db 81 -----NIDWATDGVETVLEA 95
QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFTNHDILDTGTETVISGASCTTNC 268
Db 96 TGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFTNHDILDTGTETVISGASCTTNC 155
QY 269 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMDILDPHRRGDLRRARAGANIVPNSTGAAK 328
Db 156 LAPMAKALHDAFGIQKGLMTTIHAYTGDQMDILDPHRRGDLRRARAGANIVPNSTGAAK 215
QY 329 AIGLVIPELNGKLDGAAQRPVPTGVSVELVTLDKNVSVDEINAAKAAASNDSPGYTED 388
Db 216 AIGLVIPELNGKLDGAAQRPVPTGVSVELVTLDKNVSVDEINAAKAAASNDSPGYTED 275
QY 389 PIVSSDIVGSVGSLEFDTATQTKMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 448
Db 276 PIVSSDIVGSVGSLEFDTATQTKMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 335

RESULT 2
G3P_STRP3
ID - G3P_STRP3
AC Q8K8M9; STANDARD; PRT; 335 AA.

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)
DE (Plasminogen-binding protein) (Plasmin receptor).
GN GAP OR PLR OR SPYM3_0201 OR SPS0207.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution";
RL Genome Res. 13:1042-1055(2003).
CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN (By similarity).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014140; AAM78808.1; -.
CC EMBL; AP005141; BAC63302.1; -.
CC InterPro; IPR000173; GAP_dhdrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; C; 1.
CC PRINTS; PR00078; G3PDHDRGNASE.
CC TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
FT INIT MET 0 0
FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
FT SEQUENCE 335 AA; 35841 MW; 4DCB76F382F6F98 CRC64;
Query Match 72.3%; Score 1646.5; DB 1; Length 335;
Best Local Similarity 79.0%; Pred. No. 2.1e-93;
Matches 332; Conservative 1; Mismatches 2; Indels 85; Gaps 1;
QY 29 VVKVGINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRFDGTVEVK 88

Db 1 VVKVGINGFGRIGRLAFRRIONIEGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEVK 60
QY 89 EGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVVEKGGFDVNGKFKIKVSAEK 148
Db 61 EGGFEVNGNFIKVSARDPE-----80
QY 149 DPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLEA 208
Db 81 -----NIDWATDGVLEIVLEA 95
QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGNDVKT VVFNTHDILDGTETVISGASCTTNC 268
Db 96 TGFFAKKEAAEKHLHTNGAKKVVITAPGNDVKT VVFNTHDILDGTETVISGASCTTNC 155
QY 269 LAPMAKALHDAFGIOKGLMTTHAYTGDQMIIDGPHRGGLRRRARAAGAAANIVPNSTGA 328
Db 156 LAPMAKALHDAFGIOKGLMTTHAYTGDQMIIDGPHRGGLRRRARAAGAAANIVPNSTGA 215
QY 329 AIGLVIPELNGKLDGAAQRPVPTGCVTELVVTLTKNVSVDENAAKAAASNDSEFGYT 388
Db 216 AIGLVIPELNGKLDGAAQRPVPTGCVTELVVTLTKNVSVDENAAKAAASNDSEFGYT 275
QY 389 PIVSSDIVGSYGSLEFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 448
Db 276 PIVSSDIVGSYGSLEFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 335

RESULT 3

G3P STREQ
ID G3P STREQ STANDARD; PRT; 335 AA.
AC Q59906;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)
DE (Plasminogen-binding protein) (Plasmin receptor).
GN GAP OR GAPC.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H46A;
RX MEDLINE=96305364; PubMed=8706717;
RA Gase K., Gase A., Schirmer H., Malke H.;
RT "Cloning, sequencing and functional overexpression of the
RT Streptococcus equisimilis H46A gapC gene encoding a
RT glyceraldehyde-3-phosphate dehydrogenase that also functions as a
RT plasmin(ogen)-binding protein. Purification and biochemical
RT characterization of the protein.";
RL Eur. J. Biochem. 239:42-51(1996).
CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X97788; CAA66377.1; -.
DR EMBL; Y12602; CAA73174.1; -.
DR FIR; S71350; S71350.
DR HSSP; P00362; 1GD1.

DR InterPro; IPR000173; GAP dhdrogenase.
DR InterPro; IPR006424; GAPDH-I.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-1; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD.
FT INIT MET 0 BY SIMILARITY.
FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
FT ACT_SITE 178 178 SIMILARITY).
FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
SQ SEQUENCE 335 AA; 35739 MW; FE7ACDFD7663E46 CRC64;
Query Match 68.2%; Score 1553.5; DB 1; Length 335;
Best Local Similarity 76.0%; Pred. No. 9.5e-88;
Matches 319; Conservative 2; Mismatches 14; Indels 85; Gaps 1;
QY 29 VVKVGINGFGRIGRLAFRRIONIEGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEVK 88
Db 1 VVKVGINGFGRIGRLAFRRIONIEGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEVK 60
QY 89 EGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVVEKGGFDVNGKFKIKVSAEK 148
Db 61 EGGFEVNGNFIKVSARDPE-----80
QY 149 DPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLEA 208
Db 81 -----NIDWATDGVLEIVLEA 95
QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGNDVKT VVFNTHDILDGTETVISGASCTTNC 268
Db 96 TGFFAKKEAAEKPLHANGAKKVVITAPGNDVKT VVFNTHDILDGTETVISGASCTTNC 155
QY 269 LAPMAKALHDAFGIOKGLMTTHAYTGDQMIIDGPHRGGLRRRARAAGAAANIVPNSTGA 328
Db 156 LAPMAKALHDAFGIOKGLMTTHAYTGDQMIIDGPHRGGLRRRARAAGAAANIVPNSTGA 215
QY 329 AIGLVIPELNGKLDGAAQRPVPTGCVTELVVTLTKNVSVDENAAKAAASNDSEFGYT 388
Db 216 AIGLVIPELNGKLDGAAQRPVPTGCVTELVVTLTKNVSVDENAAKAAASNDSEFGYT 275
QY 389 PIVSSDIVGSYGSLEFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 448
Db 276 PIVSSDIVGSYGSLEFDTATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 335
RESULT 4
G3P LACLA
ID G3P LACLA STANDARD; PRT; 337 AA.
AC P52987;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR LLO559.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LM0230;
RX MEDLINE=95291425; PubMed=7773380;
RA Cancilla M.R., Hillier A.J., Davidson B.E.;
RT "Lactococcus lactis glyceraldehyde-3-phosphate dehydrogenase gene,
RT gap: further evidence for strongly biased codon usage in glycolytic
RT pathway genes.";
RL Microbiology 141:1027-1036(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;

RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
lactis ssp. lactis IL1403";
RL Genome Res. 11:731-753(2001).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L36907; AAC41453.1; -.
CC EMBL; AE006290; AAK04657.1; -.
CC PIR; G86694; G86694.
CC HSSP; P17721; 1HDG.
CC InterPro; IPR000173; GAP dhdrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh C; 1.
CC PRINTS; PR00078; G3PDHDRGNASE.
CC TIGRFAMS; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; FALSE NEG.
CC Glycolysis; Oxidoreductase; NAD; Complete proteome.
KW BINDING 152 152 GLYCERALDEHYDE 3-PHOSPHATE
FT (BY SIMILARITY).
FT ACT_SITE 179 179 ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
FT CONFLICT 143 143 T -> S (IN REF. 1).
SQ SEQUENCE 337 AA; 36057 MW; 17BB8C6AAEFF589D CRC64;

Query Match 56.2%; Score 1280; DB 1; Length 337;
Best Local Similarity 61.4%; Pred. No. 4.2e-71;
Matches 259; Conservative 31; Mismatches 46; Indels 86; Gaps 2;

QY 28 MVKVGINGFGRIGRLAFRRIONVEGEVTRINDLTPNMLAHLKYDTTQGRFDGTV 87
DB 1 MVKVGINGFGRIGRLALRRIOVEGEVVAHINDLTPAMLAHLKYDTTQGRFKGTV 60

QY 88 KEGGFVNGNFTKVSARDPENIDWATDGV EIVLEALEGTV EVDKGGFDVNGKFIKVS 147
DB 61 KEDGFDVNGKFKVTAERNPE----- 81

QY 148 KDPEQIDWATDGV EIVLEIDGTV EVDKGGFDVNGKFIKVSARDPENIDWATDGV EIVLE 207
DB -----DIQWADSGVEIVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKT VVFNTHNHDILDTGTTETVISGASCTTN 267
DB 96 ATGFFATKEAEKHLHPGGAKKVLITAPGGNDVKT VVFNTHNHTILDGTTETVISAGSCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
DB 156 SLAPMADALNKNFGVKGTMTTVHSYTGDMITLDGPHRGDFFRRAAAANIVPASSGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTVTELVVTLTKNVSVDEINAAKAAASNDTSFGYTE 387
DB 216 KAIGLVLPESGLMKHQAQRVSTPTGTSITELTVLEKHVTVD EINEAMKAAADESFGYNV 275

QY 388 DPIVSSDIVGSYSLFDATQTKMEV-DGSQLVKWSWYDNEMSYTAQLVRTLEYFAKI 446
DB 276 DEIVSSDIIGMAYGSLFDATLTVTDLKDGGQLVKTAAWYDNEMSFQAQLRTLEYFAKI 335

447 AK 448

DB 336 AK 337

RESULT 5
G3P_CLOPA STANDARD; PRT; 334 AA.
AC Q59309;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) (CP
DE 17/CP 18).
GN GAP.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE FROM N.A.
RA Oster T., Assobhei O., Scherrer S., Branlant G., Branlant C.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP PARTIAL SEQUENCE OF 1-26.
RC STRAIN=W5;
RX MEDLINE=98291870; PubMed=9629918;
RA Flengsrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X72219; CAA51020.1; -.
CC PIR; S34254; S34254.
CC HSSP; P00362; 1GD1.
CC InterPro; IPR000173; GAP dhdrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh C; 1.
CC PRINTS; PR00078; G3PDHDRGNASE.
CC TIGRFAMS; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD.
KW BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 334 AA; 36078 MW; D15905D0DA7F62E7 CRC64;

Query Match 54.2%; Score 1234; DB 1; Length 334;
Best Local Similarity 59.3%; Pred. No. 2.6e-68;
Matches 249; Conservative 34; Mismatches 51; Indels 86; Gaps 3;

QY 29 VVKVGINGFGRIGRLAFRRIONVEGEVTRINDLTPNMLAHLKYDTTQGRFDGTV 88
DB 1 MTKVAINGFGRIGRLALRRILEVPGLEVVAINDLTDKMLAHLFKYDSSQGRFNGEIEVK 60

QY 89 EGGFEVNGNFIKVSARDPENIDWATDGV EIVLEALEGTV EVDKGGFDVNGKFIKVS 148
DB 61 EGAFVWNGKEVKVFAEADPEKLPW----- 84

QY 149 DPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLEA 208
Db 85 -----GEL-----GIDWVLEC 95
QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDTGTETVSGASCTTNC 268
Db 96 TGFFTKKEKAEAHVRA-GAKKVVISAPAGNDLKTIVNVNNDLDGTETVSGASCTTNC 154
QY 269 LAPMAKALHDAFGIQGLMTTHAYTGDQMLDGPGRGDLRRRAGAANIVPNSTGA 328
Db 155 LAPMAKVLNDKFGIEKGFWTTHAYTNDQNTLDGPHRKGDFRRARAAVSIIPNSTGA 214
QY 329 AIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVTLKXVSVDEINAAKASNDSEFGYTED 388
Db 215 AIAQVIPLEKGLDGNARQRPVPTGVTGLVTLVTLKXVSVDEINAAKASNDSEFGYTED 274
QY 389 PIVSSDIVGSYGSFLDQTKVMEVDGSQLVKVSVYDNEMSYTAQLVRLTLEYFAKIAK 448
Db 275 EIVSADVVGISYGSFLDQTKVMEVDGSQLVKVSVYDNEMSYTAQLVRLTLEYFAKIAK 334

RESULT 6
G3P_CLOAB STANDARD; PRT; 334 AA.
AC 052631;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR GAPC OR CAC0709.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=99392446; PubMed=10463150;
RA Schreiber W., Durre P.;
RT "The glyceraldehyde-3-phosphate dehydrogenase of Clostridium
acetobutylicum: isolation and purification of the enzyme, and
sequencing and localization of the gap gene within a cluster of other
glycolytic genes."
RL Microbiology 145:1839-1847(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF043386; AAC13160.1; -;
CC EMBL; AE007586; AAK78686.1; -;

PIR; C96987; C96987.
DR HSSP; P17721; 1HDG.
DR InterPro; IPR000173; GAP dhhydrogenase.
DR InterPro; IPR006424; GAPDH-I.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 334 AA; 35850 MW; 10C52A174BE789B5 CRC64;
Query Match 52.8%; Score 1203; DB 1; Length 334;
Best Local Similarity 57.4%; Pred. No. 2e-66;
Matches 241; Conservative 41; Mismatches 52; Indels 86; Gaps 3;
QY 29 VVKVINGFGRIQLAFRRIONVEGVETRIINDLTPNNMLAHLKYDTTQGRFDGTVEVK 88
Db 1 MAKIANGFGRIQLALRRILEVPLEVVAINDLTDAKMLAHLFKYDSSQGRFNGEIEVK 60
QY 89 EGGFEVNGNFIKVSARDPENIDWATDGVLEIVLEALEGTVEVKDGGFDVNGKFIKVSAAK 148
Db 61 EGAFVNGKEVKVFAEADPEKLPWG-----DL----- 87
QY 149 DPEQIDWATDGVLEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEIVLEA 208
Db 88 -----GIDWVLEC 95
QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDTGTETVSGASCTTNC 268
Db 96 TGFFTKKEKAEAHVRA-GAKKVVISAPAGNDLKTIVNVNNDLDGTETVSGASCTTNC 154
QY 269 LAPMAKALHDAFGIQGLMTTHAYTGDQMLDGPGRGDLRRRAGAANIVPNSTGA 328
Db 155 LAPMAKVLNDKFGIEKGFWTTHAYTNDQNTLDGPHRKGDLRRARAAVSIIPNSTGA 214
QY 329 AIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVTLKXVSVDEINAAKASNDSEFGYTED 388
Db 215 AISQVTPDLAGKLDGNAQRPVPTGVTGLVTLVTLKXVSVDEINAAKASNDSEFGYTED 274
QY 389 PIVSSDIVGSYGSFLDQTKVMEVDGSQLVKVSVYDNEMSYTAQLVRLTLEYFAKIAK 448
Db 275 PIVSADVVGINYSGLFDATLTKIVDVNGSQLVKVTAAYDNEMSYTSQLVRLTLEYFAKIAK 334

RESULT 7
G3P1_STAEF STANDARD; PRT; 336 AA.
AC 08CPY5;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH 1).
GN GAP1 OR GAP OR SE0557.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593(2003).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

Query Match 49.7%; Score 1132.5; DB 1; Length 336;
Best Local Similarity 54.7%; Pred. No. 3.9e-62;
Matches 231; Conservative 44; Mismatches 60; Indels 87; Gaps 4;

QY 28 MVMKVGINGFGRIGRLAFRRIONVEGVETPRINDLTPNNMLHLKYDTTQGRFDGTEV 87
DB 1 MAVKVAINGFGRIGRLAFRRIOVEGLEVVAVNDLTDMMHLHLKYDTTQGRFTGEV 60
QY 88 KEGGFVNGNFIVKSAERDPENIDWATDGVVEIVLEALEGTVEVKGDFVNGKFIKVS 147
DB 61 VDGGRVNGKEVKSFPDASKLPWK-----DLN----- 89
QY 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV 207
DB 90 -----IDVLE 95
QY 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTWVFNTNHDILDGTETVISG 267
DB 96 CTGYTDDKKAQHTEA-GAKKVLISAPATGDLKTIIVFNTNHOELDSETVWSGASCT 154
QY 268 CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPN 327
DB 155 SLAPVAKVLNDDFGLVEGLMTTTHAYTGDQNTQDAPHRKGDKRRARAAAEIIPNST 214
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTTELVTYLDK-NVSVDEINAAKMAASND 386
DB 215 KAIGKVIPEIDGKLDGGAQRPVATGSLTTLTVLEKQDVTVEQVNEAMKVASNES 274
QY 387 EDPVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTYEYFA 446
DB 275 EDEIVSSDVVGWMTYGSFLFDATQTRVMSVGRQLVKVAAWYDNEMSYTAQLVRL 334
QY 447 AK 448
DB 335 SK 336

RESULT 9
G3P_LACDE STANDARD; PRT; 338 AA.
ID G3P_LACDE
AC O32755;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP.
OS Lactobacillus delbrueckii (subsp. bulgaricus).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1585;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B107;
RX MEDLINE=98240227; PubMed=9579064;
RA Branny P., Delatorre F., Garel J.R.;
RT "An operon encoding three glycolytic enzymes in Lactobacillus
delbrueckii subsp. bulgaricus: glyceraldehyde-3-phosphate
dehydrogenase, phosphoglycerate kinase and triosephosphate
isomerase.";
RL Microbiology 144:905-914(1998).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; AJ000339; CAA04014.1; -
PIR; T09633; T09633.
HSSP; P17721; 1HDG.
InterPro; IPR000173; GAP_dhdrogenase.
InterPro; IPR006424; GAPDH-I.
Pfam; PF00044; gpdh; 1.
Pfam; PF02800; gpdh_C; 1.
PRINTS; PR00078; G3PDHGRNASE.
TIGRFAMS; TIGR01534; GAPDH-I; 1.
PROSITE; PS00071; GAPDH; FALSE_NEG.
KW Glycolysis; NAD; Oxidoreductase.
FT BINDING 156 156 GLYCERALDEHYDE 3-PHOSPHATE
(BY SIMILARITY).
FT ACT_SITE 183 183 ACTIVATES THIOL GROUP DURING CATALYSIS
(BY SIMILARITY).
FT SEQUENCE 338 AA; 36564 MW; DF97E8C9CA4F7DFA CRC64;

Query Match 41.9%; Score 955.5; DB 1; Length 338;
Best Local Similarity 47.3%; Pred. No. 2.3e-51;
Matches 200; Conservative 46; Mismatches 88; Indels 89; Gaps 4;
QY 28 MVMKVGINGFGRIGRLAFRRIONVEGVETPRINDLTPNNMLHLKYDTTQGRFDG 83
DB 1 MTVKIGINGFGRIGRLAFRRIMDLGEETKQIEVVAINDLTPAMLAHLKYDTHGTDFD 60
QY 84 TVEVKEGGFEVNGNFIVKSAERDPENIDWATDGVVEIVLEALEGTVEVKGDFVNGKFIK 143
DB 61 EVSATEDSLVVDGKKYVVAEPQAQNIW-----VKN----- 92
QY 144 VSAEKDPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATD 203
DB 93 -----DGVD 96

QY 204 IVLEATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTWVFNTNHDILDGTETVISG 263
DB 97 FVLECTGYTSKAKSOAHLDA-GAKRVLISAPAGNDLKTIVSVNQDTLTADDTIVSAGS 155
QY 264 CTNCLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPN 323
DB 156 CTNSLAPMANALNKEFGIQVGTMTTTHAYTATQKVLDPDRGNFRNARAAAEIIPHS 215
QY 324 TGAAKAIGLVIPELNGKLDGAAQRPVPTGSGVTTELVTYLDKNSVDEINAAKMAASND 383
DB 216 TGAAKAIGLVLPELNGKLDGHAQRPVVKDGGSETELVTILDKKVTAEVNAAMKKYESP 275
QY 384 GYTEDPIVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTYEY 443
DB 276 AYNADQIVSTDVLGMTAGSIFDPTQTQVITAGDKQLVKTVAWYDNEYSFTCQMVRL 335

444 AKI 446
336 ATL 338

RESULT 10
G3P3_ECO57 STANDARD; PRT; 333 AA.
ID G3P3_ECO57
AC P58072;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase C (EC 1.2.1.12) (GAPDH-C).
GN GAPC OR Z2304 OR ECS2022.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22 (2001).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005364; AAC56359.1; -;
DR EMBL; AP002557; BAB35445.1; -;
DR PIR; C85737; C85737.
DR PIR; F90881; F90881.
DR HSSP; P17721; 1HDG.
DR InterPro; IPR000173; GAP dhdrogenase.
DR InterPro; IPR006424; GAPDH-I.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
FT BINDING 150 150
FT (BY SIMILARITY).
FT ACT_SITE 177 177
FT ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
FT SEQUENCE 333 AA; 35763 MW; A2F77CB2E773E64C CRC64;
SQ
Query Match 41.3%; Score 941.5; DB 1; Length 333;
Best Local Similarity 47.5%; Pred. No. 1.6e-50;
Matches 199; Conservative 53; Mismatches 76; Indels 91; Gaps 7;
QY 31 KVGINGFGRIGRLAFRIQNVG-GVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEVKE 89
Db 3 KVGINGFGRIGRLVRLLEVKSNIDVVAINDLTPSKILAYLLKHDSNYGPF----- 54
QY 90 GGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAEKD 149
Db 55 ----- 54
QY 150 PEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLEAT 209
Db 55 ----PWSVDFTEDSLVD-----GKSIAYVAEKEAKNIPWAKAKAEIIVECT 97
QY 210 SFFAKKEAAEKHLHANGAKKVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTNCL 269

Db 98 GFYTSAEKSQLHDA-GAKKVLISAPAG-EMKTIIVYKVNDDTLGNDTIVSVASCITNCL 155
QY 270 APMKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAGA 329
Db 156 APMKALHDSFGIEVGTMTTHAYTGTQSLVDGP-RGKDLRASRAAENIIPHTTGAAGA 214
QY 330 IGLVIPELNGKLDGAAQRPVPTGVTTELVTLDKNVSVDEINAAKAA--SNDSFGYTE 387
Db 215 IGLVIPELSGKLKGHAQRPVPTGVTTELVSILGKVTABEVNNAKQATTNNESFGYTD 274
QY 388 DPIVSSDIVGSYSGSLFDATQTKMEVDGSQLVKVSVWVDNEMSHTAQLVRLTLEYFAKI 446
Db 275 EEIVSSDIIGSHFGSVFDTATQTEITAVGDLQLVKTVAWVDNEYGFVTQLIRLTLEKFAKL 333
RESULT 11
G3P3_ECOLI
ID G3P3_ECOLI STANDARD; PRT; 333 AA.
AC P33898; P76094; P78062; P78291; Q03850; Q63208;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase C (EC 1.2.1.12) (GAPDH-C).
GN GAPC OR B1416/B1417.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=River isolate, and Clinical isolate;
RX MEDLINE=98283700; PubMed=9622357;
RA Espinosa-Urgel M., Kolter R.;
RT "Escherichia coli genes expressed preferentially in an aquatic
RT environment";
RL Mol. Microbiol. 28:325-332 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Hidalgo E., Limon A., Aguilar J.;
RT "A second Escherichia coli gene with similarity to gapA";
RL Microbiologia 12:99-106 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474 (1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map";
RL DNA Res. 3:363-377 (1996).
RN [5]
RP SEQUENCE OF 306-332 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92011371; PubMed=1917845;
RA Hidalgo E., Chen Y.-M., Lin E.C.C., Aguilar J.;
RT "Molecular cloning and DNA sequencing of the Escherichia coli K-12
RT ald gene encoding aldehyde dehydrogenase";

J. Bacteriol. 173:6118-6123 (1991).
[6]
SEQUENCE OF 1-13 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=88232416; PubMed=2836696;
RA Nakamura H., Murakami H., Yamato I., Anraku Y.;
RT "Nucleotide sequence of the cybB gene encoding cytochrome b561 in
RT Escherichia coli K12.";
RL Mol. Gen. Genet. 212:1-5 (1988).
RN [7]
SEQUENCE OF 1-165 FROM N.A.
RP Krawetz S.A.;
RA Submitted (OCT-1990) to the EMBL/GenBank/DBJ databases.
RL [8]
SEQUENCE OF 1-52 FROM N.A.
RP MEDLINE=87053179; PubMed=3780374;
RX Krawetz S.A., Connor W., Cannon P.D., Dixon G.H.;
RA "A vector-primer-cloner-sequencer plasmid for the construction of cDNA
RT libraries: evidence for a rat glyceraldehyde-3-phosphate
RT dehydrogenase-like mRNA and a ferritin mRNA within testis.";
RL DNA 5:427-435 (1986).
RN [9]
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -!- CAUTION: IN THE K12 STRAIN THIS GENE IS DISRUPTED BY A STOP CODON
CC AND A FRAMESHIFT. IT SEEMS TO BE INTACT IN A NUMBER OF WILD
CC STRAINS.
CC -!- CAUTION: Ref.7 and Ref.8 sequences were originally thought to
CC originate from rat.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L09067; AAC23856.1; ALT_FRAME.
CC EMBL; AE000239; AAC74498.1; ALT_FRAME.
CC EMBL; AE000239; AAC74499.1; ALT_FRAME.
CC EMBL; D90780; BAA15033.1; ALT_FRAME.
CC EMBL; D90781; BAA15038.1; ALT_FRAME.
CC EMBL; M64541; -; NOT_ANNOTATED_CDS.
CC EMBL; X07569; -; NOT_ANNOTATED_CDS.
CC EMBL; X54798; CAA38569.1; -.
CC EMBL; M14166; AAA1178.1; -.
CC HSSP; P17721; 1HDG.
CC EcoGene; EG12103; gapC.
CC InterPro; IPR000173; GAP_dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh C; 1.
CC PRINTS; PR00078; G3PDHGRGNASE.
CC TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE
FT ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
FT CONFLICT 39 39 K -> Y (IN REF. 7 AND 8).
FT SEQUENCE 333 AA; 35649 MW; E88223297376B0A0 CRC64;
Query Match 41.3%; Score 940.5; DB 1; Length 333;
Best local Similarity 47.5%; Pred. No. 1.9e-50;
Matches 199; Conservative 53; Mismatches 76; Indels 91; Gaps 7;
QY 31 KVGINGFGRIGRLAFRRIONVE-GVEVTRINDLTDPNMLAHLKDYDTTQGRFDGTEVEVKE 89

Db 3 KVGINGFGRIGRLVLRLLVKSNDIDVAINDLTSPKILAYLLKHDSNYGPF----- 54
QY 90 CGFEVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVEVKDGGFDVNGKFIKVSAEKD 149
Db 55 ----- 54
QY 150 PEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLEAT 209
Db 55 -----PWSVDFTEDSLVD-----GKSIAYVAEKEAKNIPWKAKGAEIIVECT 97
QY 210 SFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVPNTNNHDLDTGTTVISGASCTTNCL 269
Db 98 GFYSAEKSAQHLDA-GAKKVLISAPAG-EMKTIYVNVNDDTLGDNDTIVSVASCTTNCL 155
QY 270 APMKALHDAFGIOKGLMTTIHAYTGDQIMLDGPHRGDLRRARAGAAANIVPNSTGAAGA 329
Db 156 APMKALHDSFGIEVGTMTTIHAYTGTQSLVDGP-PGKDLRASRAAAENIIPHTTGAAGA 214
QY 330 IGLVPELNGKLDGAAQRPVPTGVTSLVTLVTLKNSVSDVDEINAAKAA--SNDSPGYTE 387
Db 215 IGLVPELSGKLKGHAQRPVPTGVTSLVTLVTLKNSVSDVDEINAAKAA--SNDSPGYTE 274
QY 388 DPIVSSDIVGSVSGSLFDATQTKMVEVDGSQLVKVSVWYDNEMSYTAQLVRLTEYPAKI 446
Db 275 EEIVSSDIIGSHFGSVFDTQTETAVGDLQVLKTVAWYDNEYGFVTLRLTEYPAKI 333
RESULT 12
G3P_MYCPN
ID G3P_MYCPN STANDARD; PRT; 337 AA.
AC P75358;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAPA OR GAP OR MPN430 OR MP411.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449 (1996).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000040; AAB96059.1; -.
CC PIR; S73737; S73737.
CC HSSP; P17721; 1HDG.
CC InterPro; IPR000173; GAP_dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh C; 1.
CC PRINTS; PR00078; G3PDHGRGNASE.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 16:25:29 ; Search time 68.7761 Seconds
(without alignments)
2055.251 Million cell updates/sec

Title: US-10-650-369-22
Perfect score: 2278
Sequence: 1 MKKINGIILLALLAVIILSAC.....EMSYTAQLVRLTYFAKIAK 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1656.5	72.7	336	2 Q83ZF7	Q83zf7 streptococc
2	1564.5	68.7	336	2 Q9ALW2	Q9alw2 streptococc
3	1564.5	68.7	336	16 Q8E3E8	Q8e3e8 streptococc
4	1564.5	68.7	336	16 Q8DXS8	Q8dxs8 streptococc
5	1559.5	68.5	336	2 Q8KVU6	Q8kvu6 streptococc
6	1544.5	67.8	336	2 Q8GCR7	Q8gcr7 streptococc
7	1535	67.4	335	16 Q97NL1	Q97nl1 streptococc
8	1535	67.4	359	16 Q8CWN6	Q8cwn6 streptococc
9	1534.5	67.4	336	2 Q8KHG1	Q8khl1 streptococc
10	1530.5	67.2	336	2 Q8KVU5	Q8kvu5 streptococc
11	1521	66.8	337	16 Q8DVV3	Q8dvv3 streptococc
12	1513.5	66.4	336	2 Q8VVB9	Q8vvb9 streptococc
13	1459.5	64.1	320	2 Q9L5X6	Q9l5x6 streptococc
14	1410	61.9	309	2 Q9AJT7	Q9ajt7 streptococc
15	1409	61.9	309	2 Q9AJT9	Q9ajt9 streptococc
16	1407	61.8	309	2 Q9AJT4	Q9ajt4 streptococc

17	1405	61.7	309	2 Q9AJT5	Q9ajt5 streptococc
18	1404	61.6	308	2 Q9AJT8	Q9ajt8 streptococc
19	1387	60.9	305	2 Q9AJT6	Q9ajt6 streptococc
20	1357.5	59.6	336	16 Q9CDH4	Q9cdh4 lactococcu
21	1282	56.3	333	16 Q833I8	Q833i8 enterococcu
22	1205.5	52.9	336	16 Q8Y4I1	Q8y4i1 listeria m
23	1204.5	52.9	336	16 Q928H9	Q928h9 clostridium
24	1175	51.6	332	16 Q8XKT9	Q8xkt9 fusobacteri
25	1164	51.1	335	16 Q8RFN9	Q8rfn9 neisseria m
26	1145	50.3	334	16 Q9JW78	Q9jw78 neisseria m
27	1138.5	50.0	336	16 Q8CPY5	Q8cpx5 staphylococ
28	1134	49.8	330	2 Q83UX3	Q83ux3 neisseria g
29	1132.5	49.7	336	16 Q9Z5C5	Q9z5c5 staphylococ
30	1132	49.7	334	16 Q9JX95	Q9jx95 neisseria m
31	1131	49.6	330	2 Q84HZ6	Q84hz6 neisseria g
32	1130	49.6	330	2 Q84HZ4	Q84hz4 neisseria g
33	1129	49.6	335	2 Q93M61	Q93m61 staphylococ
34	1127	49.5	330	2 Q84HZ5	Q84hz5 neisseria g
35	1089	47.8	310	2 Q8L2P7	Q8l2p7 staphylococ
36	1082.5	47.5	311	2 Q8L2R0	Q8l2r0 staphylococ
37	1075	47.2	310	2 Q8L2Q2	Q8l2q2 staphylococ
38	1074	47.1	310	2 Q8L2Q6	Q8l2q6 staphylococ
39	1073	47.1	310	2 Q8L2P6	Q8l2p6 staphylococ
40	1071	47.0	310	2 Q8L2P5	Q8l2p5 staphylococ
41	1070	47.0	311	2 Q8VM79	Q8vm79 rhodospiril
42	1069	46.9	310	2 Q8L2Q3	Q8l2q3 staphylococ
43	1068.5	46.9	311	2 Q8L2P8	Q8l2p8 staphylococ
44	1067	46.8	310	2 Q8L2Q7	Q8l2q7 staphylococ
45	1065	46.8	310	2 Q8L2R7	Q8l2r7 staphylococ

ALIGNMENTS

RESULT 1
Q83ZF7 PRELIMINARY; PRT; 336 AA.
AC Q83ZF7;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Glyceraldehyde 3-P dehydrogenase.
GN GAPC.
OS Streptococcus dysgalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1334;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Casal J.F., Bolton A., Song X.-M., Willson P., Potter A.A.;
RT "Use of the surface proteins GapC and Mig of Streptococcus
RT dysgalactiae as protective antigens against mastitis in non-lactating
RT cows.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375662; AAP31408.1; -;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR00173; GAP_dhdnase.
DR InterPro; IPR00173; GAP_dhdnase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
SQ SEQUENCE 336 AA; 35928 MW; 11828218CF037076 CRC64;

Query Match 72.7%; Score 1656.5; DB 2; Length 336;
Best Local Similarity 79.6%; Pred. No. 2.3e-88;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRIQNVGVETTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 87
DB 1 MVVKVINGFGRIGRLAFRIQNVGVETTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 60

QY 88 KEGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
Db 61 KEGFEVNGNFIKVSARDPE-----81
QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 82 -----NIDWATDGEIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVLDKNVSVDEINAAKKAASNDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVLDKNVSVDEINAAKKAASNDSFGYTE 275
QY 388 DPIVSSDIVGVSGLFDTATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVGVSGLFDTATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 2
Q9ALW2 ID Q9ALW2 PRELIMINARY; PRT; 336 AA.
AC Q9ALW2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J48;
RA Seifert K.N., Bleiweis A.S., McArthur W.P., Brady L.J.;
RT "The Group B Streptococcal Surface Antigen Delta is a Glyceraldehyde 3-Phosphate Dehydrogenase."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -! SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.
DR EMBL; AF338416; AAK14387.1; -.
DR HSSP; P00362; 1GD1.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHNRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 68.7%; Score 1564.5; DB 2; Length 336;
Best Local Similarity 74.1%; Pred. No. 5e-83;
Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;
QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETVRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVKVGINGFGRIGRLAFRRIONVEGVETVRLNDLTPNMLAHLKYDTTQGRF-----54

QY 88 KEGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
Db 55 -----54
QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVLDKNVSVDEINAAKKAASNDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVATLEKDVTEEVAAMKAAANDSYGYTE 275
QY 388 DPIVSSDIVGVSGLFDTATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVGVSGLFDTATQTKVQTVTDGNQLVKVSVWYDNEMSYTSQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 3
Q8E3E8 ID Q8E3E8 PRELIMINARY; PRT; 336 AA.
AC Q8E3E8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
GN GBS1811.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Prangeul L., Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P., Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing invasive neonatal disease."
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL; AL766853; CAD47470.1; -.
DR Sagalists; gbs1811; -.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHNRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Complete proteome.
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 68.7%; Score 1564.5; DB 16; Length 336;
Best Local Similarity 74.1%; Pred. No. 5e-83;
Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;
QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETVRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVKVGINGFGRIGRLAFRRIONVEGVETVRLNDLTPNMLAHLKYDTTQGRF-----54

Db 1 MVVKVINGFGRIGRLAFRRRIQNVGEVETVTRINDLTDPNMLAHLKYDTTQGRF----- 54
QY 88 KEGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSARE 147
Db 55 ----- 54
QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 96 ATGFFASKEKAEQHIHENGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGGPHRGDLRRRAGAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGGPHRGDLRRRAGAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKKAASNDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKKAASNDSFGYTE 275
QY 388 DPIVSSDIVSGSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVSGSLFDATQTKVQTVDGNQLVKVSVWYDNEMSYTSQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336
RESULT 4
Q8DXS8 PRELIMINARY; PRT; 336 AA.
ID Q8DXS8
AC Q8DXS8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
GN GAP OR SAG1768.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,
Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL; AE014272; AAN00631.1; -;
DR TIGR; SAG1768; -;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Complete proteome.
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 68.7%; Score 1564.5; DB 16; Length 336;
Best Local Similarity 74.1%; Pred. No. 5e-83;
Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRRRIQNVGEVETVTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRRIQNVGEVETVTRINDLTDPNMLAHLKYDTTQGRF----- 54
QY 88 KEGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSARE 147
Db 55 ----- 54
QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 55 -----DGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 96 ATGFFASKEKAEQHIHENGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGGPHRGDLRRRAGAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGGPHRGDLRRRAGAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKKAASNDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKKAASNDSFGYTE 275
QY 388 DPIVSSDIVSGSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVSGSLFDATQTKVQTVDGNQLVKVSVWYDNEMSYTSQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336
RESULT 5
Q8KVU6 PRELIMINARY; PRT; 336 AA.
ID Q8KVU6
AC Q8KVU6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
OS Streptococcus uberis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1349;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 9927;
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
Potter A.A.;
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
antigens confers protection against heterologous challenge with
Streptococcus uberis.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF421900; AAM73771.1; -;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.

```
SQ SEQUENCE 336 AA; 35906 MW; 755D74A4548E61D4 CRC64;
Query Match 68.5%; Score 1559.5; DB 2; Length 336;
Best Local Similarity 73.9%; Pred. No. 9.7e-83;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRRIQNVEGVETTRINDLTDPMNLALHLLKYDTTQGRFDGTV 87
Db 1 MVVKVINGFGRIGRLAFRRRIQNVEGVETTRINDLTDPMNLALHLLKYDTTQGRFDGTV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGV EIVLEALEGTV EVDKGGFDVNGKFIKVS 147
Db 61 KGGFVNGNFIKVSARDPENIDWATDGV EIVLEALEGTV EVDKGGFDVNGKFIKVS 81

QY 148 KOPEQIDWATDGV EIVLEIDGTV EVDKGGFVNGQFVKVSAEREPANIDWATDGV EIVLE 207
Db 82 -----NIDWATDGV EIVLE 95

QY 208 ATFFFAKKEAAEKHLHANGAKKVVITAPGGNDVKT VVFNTHDILDGTETVISGASCTTN 267
Db 96 ATFFFAKKEAAEKHLHANGAKKVVITAPGGDDVKT VVFNTHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDDQMLDGP HRRRAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDDQMLDGP HRRRAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSVDEINAAKAAANDSF 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVA VLEKETSVEEINAAKAAANDSF 275

QY 388 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIIGMAYGSLFDATQTKVQTV DGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448
Db 336 K 336

RESULT 6
Q8GCR7 Q8GCR7 PRELIMINARY; PRT; 336 AA.
AC Q8GCR7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12).
OS Streptococcus suis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1307;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S735;
RA Brassard J., Gottschalk M., Quessey S.;
RT "Cloning and purification of Streptococcus suis serotype 2
glyceraldehyde-3-phosphate dehydrogenase.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY167026; AAN86058.1;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR00173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 35825 MW; 1D8B9B1A492DCF59 CRC64;
Query Match 67.8%; Score 1544.5; DB 2; Length 336;
```

```
Best Local Similarity 72.9%; Pred. No. 7.2e-82;
Matches 307; Conservative 16; Mismatches 13; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRRIQNVEGVETTRINDLTDPMNLALHLLKYDTTQGRFDGTV 87
Db 1 MVVKVINGFGRIGRLAFRRRIQNVEGVETTRINDLTDPMNLALHLLKYDTTQGRF----- 54

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGV EIVLEALEGTV EVDKGGFDVNGKFIKVS 147
Db 55 ----- 54

QY 148 KOPEQIDWATDGV EIVLEIDGTV EVDKGGFVNGQFVKVSAEREPANIDWATDGV EIVLE 207
Db 55 -----DGV EVDKGGFVNGKFKVSAEREPGNIDWATDGV DIVE 95

QY 208 ATFFFAKKEAAEKHLHANGAKKVVITAPGGNDVKT VVFNTHDILDGTETVISGASCTTN 267
Db 96 ATFFFAKKEAAEQHIHANGAKKVVITAPGGNDVKT VVFNTHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDDQMLDGP HRRRAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALHDAFGVQKGLMTTHIGYTGDDQMLDGP HRRRAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSVDEINAAKAAANDSF 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVA TLDKKTAAEVNAAKAAATESF 275

QY 388 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DQLVSSDIVGVSGLFDATQTKVIEVDGEQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448
Db 336 K 336

RESULT 7
Q97NLI Q97NLI PRELIMINARY; PRT; 335 AA.
AC Q97NLI;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12).
GN SP2012 OR GAPDH.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
pneumoniae.";
RL Science 293:498-506 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC11733;
RA Bergmann S., Hammerschmidt S.;
RT "Identification of pneumococcal GAPDH as plasmin(ogen)-binding
protein.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE007490; AAK76079.1;
DR EMBL; AJ505822; CAD44376.1; --
```

McAhren S.M., McHenney M., McLeaster K., Mundy C.W., Nicas T.I., Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P., Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G., Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L., Glass J.I.,
"Genome of the bacterium Streptococcus pneumoniae strain R6.";
J. Bacteriol. 183:5709-5717(2001).
EMBL; AE008547; AAL00628.1; -
PIR; G98099; G98099.
GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
GO; GO:0016491; F:oxidoreductase activity; IEA.
GO; GO:0006096; P:glycolysis; IEA.
InterPro; IPR006424; GAPDH-I.
InterPro; IPR000173; GAP_dhdrogenase.
Pfam; PF00044; gpdh; 1.
Pfam; PF02800; gpdh_C; 1.
PRINTS; PR00078; G3PDHDRGNASE.
TIGRFAMs; TIGR01534; GAPDH-I; 1.
PROSITE; PS00071; GAPDH; 1.
Oxidoreductase; Complete proteome.
SEQUENCE 335 AA; 35856 MW; DA483CEA423E747B CRC64;
Query Match 67.4%; Score 1535; DB 16; Length 335;
Best Local Similarity 73.6%; Pred. No. 2.5e-81;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
Db 61 KEGGFVNGKFIKVSARDPE----- 81
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVLEIVLE 207
Db 82 -----QIDWATDGVLEIVLE 95
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTGTEVISGASCTTN 267
Db 96 ATGFPAKKEAAEKHL-KGGAKKVVITAPGGNDVKTVVFNTHDILDTGTEVISGASCTTN 154
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAA 327
Db 155 CLAPMAKALQDNFVGVVEGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAA 214
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLTKNVSVDENAAKKAASNDSEGYTE 387
Db 215 KAIGLVIPELNGKLDGSAQRPVPTGTSVTELVAVLEKNVTVDENAAKKAASNESYGYTE 274
QY 388 DPIVSSDIVGVSYSLSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 275 DPIVSSDIVGMSYLSLFDATQTKVLDVKGQLVKVSWSYDNEMSYTAQLVRLTLEYFAKIA 334
QY 448 K 448
Db 335 K 335
RESULT 8
Q8CWN6 PRELIMINARY; PRT; 359 AA.
AC Q8CWN6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (Phosphorylating)
DE (EC 1.2.1.12).
GN GAPA OR SPR1825.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=21429245; PubMed=11544234;
RX Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Bargett S., DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C., Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E., LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,

McAhren S.M., McHenney M., McLeaster K., Mundy C.W., Nicas T.I., Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P., Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G., Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L., Glass J.I.,
"Genome of the bacterium Streptococcus pneumoniae strain R6.";
J. Bacteriol. 183:5709-5717(2001).
EMBL; AE008547; AAL00628.1; -
PIR; G98099; G98099.
GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
GO; GO:0016491; F:oxidoreductase activity; IEA.
GO; GO:0006096; P:glycolysis; IEA.
InterPro; IPR006424; GAPDH-I.
InterPro; IPR000173; GAP_dhdrogenase.
Pfam; PF00044; gpdh; 1.
Pfam; PF02800; gpdh_C; 1.
PRINTS; PR00078; G3PDHDRGNASE.
TIGRFAMs; TIGR01534; GAPDH-I; 1.
PROSITE; PS00071; GAPDH; 1.
Oxidoreductase; Complete proteome.
SEQUENCE 335 AA; 35856 MW; DA483CEA423E747B CRC64;
Query Match 67.4%; Score 1535; DB 16; Length 335;
Best Local Similarity 73.6%; Pred. No. 2.5e-81;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
Db 61 KEGGFVNGKFIKVSARDPE----- 81
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVLEIVLE 207
Db 82 -----QIDWATDGVLEIVLE 95
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTGTEVISGASCTTN 267
Db 96 ATGFPAKKEAAEKHL-KGGAKKVVITAPGGNDVKTVVFNTHDILDTGTEVISGASCTTN 154
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAA 327
Db 155 CLAPMAKALQDNFVGVVEGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAA 214
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTSVTELVTLTKNVSVDENAAKKAASNDSEGYTE 387
Db 215 KAIGLVIPELNGKLDGSAQRPVPTGTSVTELVAVLEKNVTVDENAAKKAASNESYGYTE 274
QY 388 DPIVSSDIVGVSYSLSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 275 DPIVSSDIVGMSYLSLFDATQTKVLDVKGQLVKVSWSYDNEMSYTAQLVRLTLEYFAKIA 334
QY 448 K 448
Db 335 K 335
RESULT 8
Q8CWN6 PRELIMINARY; PRT; 359 AA.
AC Q8CWN6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (Phosphorylating)
DE (EC 1.2.1.12).
GN GAPA OR SPR1825.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=21429245; PubMed=11544234;
RX Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Bargett S., DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C., Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E., LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,

RP SEQUENCE FROM N.A.
RC SPECIES=S.agalactiae, and S.iniae; STRAIN=ATCC 27541, and 9117;
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
RA Potter A.A.;
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT antigens confers protection against heterologous challenge with
RT Streptococcus uberis.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF421899; AAM73770.1; --
DR EMBL; AF421902; AAM73773.1; --
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 35723 MW; ABAB9E14F3ED1111 CRC64;

Query Match 67.4%; Score 1534.5; DB 2; Length 336;
Best Local Similarity 72.7%; Pred. No. 2.7e-81;
Matches 306; Conservative 17; Mismatches 13; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLPNMLAHLKYDTTQGRFDGTV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLPNMLAHLKYDTTQGRFDGTV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 61 KDGFEVNGS----- 70

QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV 207
Db 71 -----FVKVSAEREPANIDWATDGV 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVITAPGGNDVKTVPNTNHDILDGTETVISCASCT 267
Db 96 ATGFFASKAAAEQHIHANGAKKVITAPGGNDVKTVPNTNHDILDGTETVISCASCT 155

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANVPNSTGA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANVPNSTGA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLTKVSVDEINAAKAAANDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLTKVSVDEINAAKAAANDSFGYTE 275

QY 388 DPVSSDIVGVSYSGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAK 447
Db 276 DAIVSSDIVGISYGLFDATQTKVQTVVDGNQLVKVSVWYDNEMSYTAQLVRLTLEYFAK 335

QY 448 K 448
Db 336 K 336

RESULT 10
Q8KVU5
ID Q8KVU5 PRELIMINARY; PRT; 336 AA.
AC Q8KVU5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
OS Streptococcus parauberis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

OX NCBI_TaxID=1348;
RN [1]
RP SEQUENCE FROM N.A.
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
RA Potter A.A.;
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT antigens confers protection against heterologous challenge with
RT Streptococcus uberis.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF421901; AAM73772.1; --
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 36037 MW; 311647C25489AC9E CRC64;

Query Match 67.2%; Score 1530.5; DB 2; Length 336;
Best Local Similarity 73.4%; Pred. No. 4.6e-81;
Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLPNMLAHLKYDTTQGRFDGTV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLPNMLAHLKYDTTQGRFDGTV 55

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 56 -----GTVEVKDGGFDVNGKFIKVS 77

QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV 207
Db 78 KDPEQ-----IDWATDGV 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVITAPGGNDVKTVPNTNHDILDGTETVISCASCT 267
Db 96 ATGFFAKKAAAEKHLHANGAKKVITAPGGDDVKTVPNTNHDILDGTETVISCASCT 155

QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANVPNSTGA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANVPNSTGA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLTKVSVDEINAAKAAANDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLTKVSVDEINAAKAAANDSFGYTE 275

QY 388 DPVSSDIVGVSYSGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAK 447
Db 276 DPVSSDIVGMSFGSLFDATQTKVQTVVDGNQLVKVSVWYDNEMSYTAQLVRLTLEYFAK 335

QY 448 K 448
Db 336 K 336

RESULT 11
Q8DVV3
ID Q8DVV3 PRELIMINARY; PRT; 337 AA.
AC Q8DVV3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Extracellular glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12).
GN GAPC OR SMU.360.
OS Streptococcus mutans.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
DR EMBL; AE014883; AAN58118.1; --
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; Spdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
DR Oxidoreductase; Complete proteome.
KW SEQUENCE 337 AA; 36068 MW; 42BFE20365963C22 CRC64;
SQ

Query Match 66.8%; Score 1521; DB 16; Length 337;
Best Local Similarity 71.8%; Pred. No. 1.7e-80;
Matches 303; Conservative 18; Mismatches 15; Indels 86; Gaps 2;
QY 28 MVVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTV 87
Db 1 MVVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDSTQGRFDGTV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 61 KEGGFVNGKFKVSAERDPE----- 81
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLE 207
Db 82 -----QIDWAADGVLE 95

208 ATSFPAKAEAEKHLHAN-GAKKVVITAPGGNDVKTVFNTNHDILDGTETVISGASCTT 266
96 ATGFFASKAAAEKHLHANGGAKKVVITAPGGNDIKTIVFNTNHDVLDGTETVISGASCTT 155
267 NCLAPMAKALHDAFGIQKGLMTTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGA 326
156 NCLAPMAKALHDNFSIKEGLMTTTHAYTGDMVLDGPHRKGLRRARAGAAANIVPNSTGA 215
327 AKAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKAAASNDSTFGYT 386
216 AKAIGLVIPELNGKLDGAAQRPVPTGVSVELVAVLTKKVTVDENAAKAAANESYGYT 275
387 EDPIVSSDIVGVSGLFDATQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKI 446
276 EDPIVSSDIVGVSGLFDATQTKVLDVDGKQLVKVSWYDNEMSYTSQLVRLTLEYFAKI 335
447 AK 448
336 AK 337

RESULT 12
Q8VVB9 PRELIMINARY; PRT; 336 AA.
ID Q8VVB9;
AC Q8VVB9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
GN GAPDH.

OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LMG18311;
RA van den Bogaard P.T.C., Kleerebezem M., Hols P., Crispie F.,
RA Kuipers O.P., de Vos W.M.;
RT "Modulation of Glycolysis by Lactose Availability in Streptococcus
RT thermophilus.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF442551; AAL35377.1; --
DR HSSP; P00354; 3GPD.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
DR Oxidoreductase.
KW SEQUENCE 336 AA; 36026 MW; 52C1F25F3A7E0230 CRC64;
SQ

Query Match 66.4%; Score 1513.5; DB 2; Length 336;
Best Local Similarity 71.7%; Pred. No. 4.5e-80;
Matches 302; Conservative 18; Mismatches 16; Indels 85; Gaps 1;
QY 28 MVVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTV 87
Db 1 MVVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVS 147
Db 61 KEGGFVNGKFKVSAERDPE----- 81
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLE 207
Db 82 -----QIDWANDGVLE 95

208 ATSFPAKAEAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDGTETVISGASCTTN 267
96 ATGFFTKKVLAEKHLHPGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTN 155
268 CLAPMAKALHDAFGIQKGLMTTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
156 CLAPMAKALNDNFGIVGELMTTTHAYTGDMILDGPNRGGDLRRARAGAAANIVPNSTGAA 215
328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDENAAKAAASNDSTFGYTE 387
216 KAIGLVIPELNGKLDGSAQRPVPTGVSVELVAVCEKNVTVDENAAKAAATNESYGYTE 275
388 DPIVSSDIVGVSGLFDATQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
276 DPIVSSDIVGVSGLFDATQTKVLDVDGKQLVKVSWYDNEMSYTSQLVRLTLEYFAKIA 335
448 K 448
336 K 336

RESULT 13
Q9L5X6 PRELIMINARY; PRT; 320 AA.
ID Q9L5X6;
AC Q9L5X6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Extracellular glyceralddehyde-3-phosphate dehydrogenase
DE (Fragment).
OS Streptococcus gordonii.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1302;
RN [1]
RP SEQUENCE FROM N.A.
RA Nelson D., Goldstein J.M., Boatright K., Harty D.W.S., Cook S.L.,
RA Hickman P.J., Potempa J., Travis J., Mayo J.A.;
RT "Purification and characterization of an extracellular
RT glyceralddehyde-3-phosphate dehydrogenase from Streptococcus sanguis
RT and cloning of the gene encoding this enzyme."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF247678; AAF64063.1; -.
DR HSSP; P00362; 1GDI.
DR GO; GO:0004365; F:glyceralddehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 1
FT NON_TER 320
FT NON_TER 320
SQ SEQUENCE 320 AA; 33968 MW; 6FFCC18BB01E91F0 CRC64;

Query Match 64.1%; Score 1459.5; DB 2; Length 320;
Best Local Similarity 71.9%; Pred. No. 5.7e-77;
Matches 291; Conservative 14; Mismatches 15; Indels 85; Gaps 1;
QY 31 KVGINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKYYDTTQGRFDGTVEVKEG 90
Db 1 KVGINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKYYDTTQGRFDGTVEVKEG 60
QY 91 GFEVNGNFIKVSARDPENIDWATDGVLEALEGTVKVGKGFVNGKFIKVSAREKDP 150
Db 61 GFEVNGKFKVSAERDPE-----NIDWANDGVLEALEATG 78
QY 151.EQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEALEATS 210
Db 79 -----NIDWANDGVLEALEATG 95
QY 211 FFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISCASCTTNCLA 270
Db 96 FFAKKAEEKHLHAGGAKKVVITAPGGSDVKTVVFNTHDVLDTGTETVISCASCTTNCLA 155
QY 271 PMAKALHDAFGIQGLMTTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAKAI 330
Db 156 PMAKALQDNFVGVVEGLMTTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAKAI 215
QY 331 GLVPELNGKLDGAAQRPVPTGVTGLVTLVTLTKNVSVDENAAKAAASNDSPGYTEDPI 390
Db 216 GLVPELNGKLDGAAQRPVPTGVTGLVTLVTLTKNVTVDENAAKAAASNESYGYTEDPI 275
QY 391 VSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQ 435
Db 276 VSSDIVGVSGLFDATQTKVLDVKGKQLVKVSWYDNEMSYTAQ 320

RESULT 14
Q9AJT7 PRELIMINARY; PRT; 309 AA.
ID Q9AJT7
AC Q9AJT7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
DE (Fragment).
GN GAP.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RA Amzaga M.R., Carter P.E., Cash P., McKenzie H.;
RT "Sequence variation in mef(A) and gap genes in M-phenotype
RT erythromycin resistant Streptococcus pneumoniae";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AJ292048; CAC27448.1; -.
DR HSSP; P00362; 1GDI.
DR GO; GO:0004365; F:glyceralddehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 1
FT NON_TER 309
FT NON_TER 309
SQ SEQUENCE 309 AA; 32930 MW; 89E45C3BEA6A0528 CRC64;

Query Match 61.9%; Score 1410; DB 2; Length 309;
Best Local Similarity 71.9%; Pred. No. 4e-74;
Matches 284; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
QY 36 GFGRIGRILAFRRIONVEGVETTRINDLTPNMLAHLKYYDTTQGRFDGTVEVKEGGFEV 95
Db 1 GFGRIGRILAFRRIONVEGVETTRINDLTPNMLAHLKYYDTTQGRFDGTVEVKEGGFEV 60
QY 96 GNFIKVSARDPENIDWATDGVLEALEGTVKVGKGFVNGKFIKVSAREKDPQIDW 155
Db 61 GNFIKVSARDPE-----QIDWATDGVLEALEATGFFAKK 73
QY 156 ATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEALEATSFFAKK 215
Db 74 -----QIDWATDGVLEALEATGFFAKK 95
QY 216 EAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISCASCTTNCLAPMAKA 275
Db 96 EAAEKHL-KGAKKVVITAPGGNDVKTVVFNTHDVLDTGTETVISCASCTTNCLAPMAKA 154
QY 276 LHDAFGIQGLMTTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAGAIGLVIP 335
Db 155 LQDNFVGVVEGLMTTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAAGAIGLVIP 214
QY 336 ELNGKLDGAAQRPVPTGVTGLVTLVTLTKNVSVDENAAKAAASNDSPGYTEDPIVSSDI 395
Db 215 ELNGKLDGSAQRPVPTGVTGLVTLVTLTKNVTVDENAAKAAASNESYGYTEDPIVSSDI 274
QY 396 VGVSYSGLFDATQTKVMEVDGSQLVKVSWYDNEM 430
Db 275 VGMSYSGLFDATQTKVLDVKGKQLVKVSWYDNEM 309

RESULT 15
Q9AJT9 PRELIMINARY; PRT; 309 AA.
ID Q9AJT9
AC Q9AJT9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
DE (Fragment).
GN GAP.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=clinical isolate;
RA Amezaga M.R., Carter P.E., Cash P., McKenzie H.;
RT "Sequence variation in mef(A) and gap genes in M-phenotype
RT erythromycin resistant Streptococcus pneumoniae.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
CC EMBL; AJ292046; CAC27446.1; --
DR HSSP; P00362; IGD1.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHHRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 1
FT NON_TER 309
SQ SEQUENCE 309 AA; 32916 MW; 39E9CC3BEA6F8959 CRC64;

Query Match 61.9%; Score 1409; DB 2; Length 309;
Best Local Similarity 71.6%; Pred. No. 4.6e-74;
Matches 283; Conservative 13; Mismatches 13; Indels 86; Gaps 2;
QY 36 GFGRIGLAFRRIONVEGEVETRINDLTPNMLAHLKYDTTQGRFDGTVEVKEGGFEVN 95
Db 1 GFGRIGLAFRRIONVEGEVETRINDLTPNMLAHLKYDTTQGRFDGTVEVKEGGFEVN 60
QY 96 GNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAREKDPEDQIDW 155
Db 61 GKFKVSAERDPE-----QIDWATDGVIEVLEATGFFAKK 73
QY 156 ATDGVIEVLEIDGTVEVKEGGFEVNGQVKVSAEREANIDWATDGVIEVLEATSPFAKK 215
Db 74 -----QIDWATDGVIEVLEATGFFAKK 95
QY 216 EAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVSGASCTTNCLAPMAKA 275
Db 96 EAAEKHL-KGGAKKVVITAPGGNDVKTIVFNTNHDVLDGTETVSGASCTTNCLAPMAKA 154
QY 276 LHDAFGIQKGLMTTIHAYTGDQMILDGPHRGDLRRARAGAANIVPNSTGAAGAIGLVIP 335
Db 155 LQDNFGVVEGLMTTIHAYTGDQMILDGPHRGDLRRARAGAANIVPNSTGAAGAIGLVIP 214
QY 336 ELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKKAASNDSPGYTEDPIVSSDI 395
Db 215 ELNGKLDGSAQRVPTGVSVELVAVLEKNVTVEVNAAMKAASNESYGYTEDPIVSSDI 274
QY 396 VGVSYSGLFDATQTKVMEVDGSQLVKVSWYDNEM 430
Db 275 VGMSYSGLFDATQTKVLDVKGKQLVKVSWYDNEM 309